AFML-TR-67-379

INDEXER PERFORMANCE ANALYSIS AND OPERATIONS OF A DOCUMENT RETRIEVAL SYSTEM

F. L. SCHEFFLER

UNIVERSITY OF DAYTON RESEARCH INSTITUTE
DAYTON, OHIO

TECHNICAL REPORT AFML-TR-67-379

FEBRUARY 1968

AIR FORCE MATERIALS LABORATORY
AIR FORCE SYSTEMS COMMAND
WRIGHT-PATTERSON AIR FORCE BASE, OHIO

84

INDEXER PERFORMANCE ANALYSIS AND OPERATIONS OF A DOCUMENT RETRIEVAL SYSTEM

F. L. SCHEFFLER

FOREWORD

This report was prepared by the University of Dayton Research Institute, Dayton, Ohio, under Air Force Contract AF 33(615)-3389. The work described herein was accomplished under Project No. 7381 "Materials Application," and Task No. 738103, "Materials Information Development, Collection and Processing." The effort was administered under the direction of the Materials Information Branch, Materials Applications Division, Air Force Materials Laboratory with H. B. Thompson, MAAM, as project monitor.

This is a summary technical report and covers the work accomplished from 1 December 1966 through 30 November 1967.

The author acknowledges the efforts and contributions of the Project Supervisor, Edward A. Janning, his co-workers Eugene R. Egan, Ralph B. Smith, and Linda C. Haley, and the student indexers, Patricia A. Thomas and Edward J. Fruin. Dr. Ivan Goldfarb from the Polymer Branch, AFML, served as a consultant on search profile preparation and Edward L. Horne of the Materials Information Branch, AFML, coordinated the CAS-SDI program.

This report was submitted by the author November 1967.

This technical report has been reviewed and is approved.

D. A. SHINN
Chief, Materials Information Branch
Materials Applications Division
Air Force Materials Laboratory

ABSTRACT

Two graduate students in chemistry were trained in indexing of technical documents through an established training program. Certain modifications were made which tended to improve the training as evidenced by improved performance. A study of intra- and interindexer consistency was made by subjecting a group of technical documents to independent indexing by various individuals and by two of the same individuals one year later. The concept of essential index terms was introduced. Statistical analysis showed that intra- and interindexer correlation for experienced indexers of essential terms is significant with a probability (p) > 0.999.

An evaluation of the indexer trainees showed that essential terms was a better performance criterion than total terms and that high interindexer consistency was achieved. Search requests declined from the preceding year, especially from the AFML. Possible reasons for the decline are presented. A Selective Dissemination of Information (SDI) Program with Chemical Abstracts was initiated. The development of a vocabulary and thesaurus for information science related documents has started. A free indexing technique is being used.

TABLE OF CONTENTS

SECTION I

INTRODUCTION

The Information Systems Section of the University of Dayton Research Institute, under contract to the Air Force Materials Laboratory (AFML), has established and presently maintains a document retrieval system in support of the Aerospace Materials Information Center (AMIC). The AFML, through its Materials Information Branch, has established a network of seven specialized information and data centers throughout the country which constitute the Air Force Materials Information Centers (AFMIC). The purpose of AFMIC is to acquire, process, store, retrieve, and disseminate materials data and information to members of the scientific and technical community engaged in research and development for the Department of Defense (DOD). The Air Force encourages the use of these centers which provide their services to Government agencies, Government contractors, suppliers and others who are in a position to contribute to the defense effort.

The document retrieval system operated by the University of Dayton contains approximately 31,000 documents related to materials research and development with new accessions being added continually. The establishment, modification, and operation of the document retrieval system are described in References 6, 7, 8, and 10. A recapitulation of these reports is presented in Appendix I.

This report covers the work accomplished from December 1966 to December 1967. The indexer training program developed and described in Reference 10 was used in indexing training of two graduate students in chemistry and a high school teacher in physics. Although the same basic program was followed, certain modifications were included which improved the effectiveness of the training. A study of the criteria for evaluating new indexers was made and some interesting results of intra- and interindexer consistency were obtained by statistical analysis. A pilot Selective Dissemination of Information (SDI) program was inaugurated with Chemical Abstracts Service (CAS). The CAS services subscribed to were the SDI experiment, POST-P, and POST-J (See Section IV). Various changes in the operations of the document retrieval system were made, and documents related to information science are being indexed for eventual inclusion in the system. The terminology of information science presents special problems for indexing and thesaurus development which will be considered before a viable retrieval system for these documents can be established. An exciting new concept called computer assisted indexing has been evolved for processing of documents. This study has just commenced. The acquisition of an IBM 360 computer at the University of Dayton in conjunction with McCall Corporation, and the scheduled acquisition of remote terminal access by the Information Systems Section will provide the hardware necessary for this study.

SECTION II

USE OF GRADUATE STUDENTS AS INDEXERS

Last year, encouraging results were obtained with a chemistry student, who had just graduated, and another student, who had just completed her junior year in chemistry, as indexers. This experience showed the feasibility of using technically trained students for indexing materials documents and suggested the employment of additional student personnel. During the summer of 1967, two graduate students in chemistry were obtained on graduate assistantships in cooperation with the Department of Chemistry. In the graduate assistantship program, the students work full time during the summer and half-time during the school year while attending classes. It is a two-year program, thus providing the services of the students for an extended period of time after they have been trained.

The two students started to work early in July 1967 and their training followed very closely that described in detail in Reference 10. About the same time, a high school physics teacher was employed for full time summer employment and part-time work during the school year. The three persons had the opportunity to proceed through the training program at the same time.

There were certain aspects of the original training program which were modified as indicated by the progress of the indexer trainees. The extracting function was de-emphasized to some extent, and direct indexing from the document was started earlier. The extracting of documents, that is, the writing out of concepts in word, phrase, or sentence form, tended to regiment one's thinking to consideration of the document as a whole. Further, extracting tended to require excessive time, and the procedure was not at all commensurate with the benefit derived. Document analysis, however, was effective in learning the indexing function. As will be shown in the next section, the average indexing time required for the series of documents used in experimental testing was remarkably less for this year's trainees. The total formal training period was reduced by 25%, from four weeks in 1966 to three weeks in 1967. Comparisons of the students from both years are presented in Section III.

It can be stated that the results of the indexer training program from 1966 were very good, and the results from 1967 were even better. There are several factors which may account for this. The good results from both years attributed to considerable interaction between the trainees and the supervisors. The trainees received rapid feedback at first by independ indexing of the same documents by the supervisors and trainees and later by review indexing of the trainees' index cards. The success of the program is also due to some extent to the exposure of the students to the entire operation

of the document retrieval system. In this way, they see the importance of the original indexing to the other functions, especially in the retrieval of documents. Also the students are periodically apprised of other Information Systems Section projects and are exposed to other information systems.

The indexers are performing professional work of basic importance to the continued operations of the document retrieval system, and they are treated as professionals. The indexers are encouraged to contribute ideas for improving operation, and numerous worthwhile suggestions have been received.

The rate of advancement of the indexer trainees was more rapid this year than last year. One reason for this was that the administration of the training program was better, since experience had been gained from applying the program the preceding year. The de-emphasizing of the extracting procedure was also helpful. Concomitant with these factors the indexers were more experienced and mature, and hence their ability to grasp the material more quickly was enhanced. Since the graduate assistants are committed to pursuing their graduate course of study at the University of Dayton for two years, they may be more highly motivated than the students who were employed only for the summer.

In conclusion, it may be stated that the indexer training program developed at the University of Dayton is highly effective in producing competent indexers. Both undergraduate and graduate students have been trained, with somewhat better results being obtained with graduate students. It cannot be stated categorically, however, that graduate students necessarily become better indexers, since other factors were operating which may as well account for this result.

SECTION III

A STUDY OF INTRA- AND INTERINDEXER CONSISTENCY AND INDEXING EVALUATION

As reported in Reference 10, an evaluation procedure has been developed to evaluate the progress of indexer trainees. In essence, the technique compared the indexing terminology used by the students with the total terms used by the supervisor in indexing and also with the essential terms used by the supervisor in indexing. The time required for indexing was also compared. It was suggested that a fairer, more objective evaluation would be made by using the essential terms as a basis for comparison rather than by using the total terms used by the supervisor. Implicit in using the supervisor's indexing as a reference standard is an assumed reliability or consistency of the supervisor himself regarding his indexing.

The use of essential terms suggests that these terms have certain characteristics which distinguish them along a continuous subjective parameter of "essentiality" so that index terms can be judged as being essential or nonessential. One could state that if an index term is listed, it must be essential or it would not be listed. However, the philosophy of indexing employed in the Information Systems Section is that while there are essential terms which must be listed, there are other terms which may well be helpful in indexing and retrieval but which do not have the character of being essential. The judgment of essential terms is tempered by experience with search requests. The frequency with which terms are used in formulating search strategies has some influence on the judgment of what constitutes essential terms. Since the AMIC document retrieval system is oriented toward materials information, materials terms are usually essential terms, and this is reflected in the user requests. For a system fulfilling a different mission, e.g., a system containing human factors information, materials terms probably would not be considered essential terms.

Assuming "ideal" indexing is not always possible, it is considered better to "overindex" a document, risking a false retrieval, than to "underindex" the document, risking nonretrieval of a pertinent item. If the document is retrieved and presented, the requestor at least has the option of accepting or rejecting it; if it is not retrieved, this selection option cannot be exercised. Generally, terms from the title, terms from the abstract, and materials terms are considered essential, but there are many exceptions to this generalization depending on the document, the subject matter, etc. Essential terms should be less subject to variation than the helpful, but not essential, terms.

Studies of intra- and interindexer consistency were undertaken to determine the validity of using an experienced indexer's indexing as a reference standard and to assess the quality of indexing among various indexers. One might consider the following question: if a document were given to an indexer for indexing, what degree of correlation of terms would result if the same document were indexed by the same indexer at a different time? Another question would be as follows: What correlation of terms would result from two independent indexers indexing the same document? If it can be shown that statistically significant correlation exists, evidence would be provided for assessing intra- and interindexer reliability or consistency. It was to the answering of these questions that the following studies were directed.

In 1966, 35 documents were selected which dealt primarily with chemistry. These documents were independently indexed by both trainees and the supervisor, an experienced indexer, about one week after the trainees had completed their training program. Several months later the experiment was repeated with another series of documents. Comparisons were made of essential terms, total terms, and time required for indexing. The student trainees were instructed to index the document as well as they could; they were unaware of the essential terms until the completion of the first experiment. The students' indexing was simply compared with the supervisor's judgment of essential and total terms as a reference standard. The results were reported in tabular form in Reference 10.

In 1967, the first set of 35 documents was recalled from the AFML Library and indexing was performed by the two new trainees and by two experienced indexers, one of whom was the supervisor in 1966. Both experienced indexers were instructed to designate the essential terms for each indexed document. Although the same documents were re-indexed, there was little noticeable learning effect, since an entire year had elapsed with many documents being indexed during the year by both experienced indexers. One indexer, however, indicated that the prior indexing had made the documents somewhat familiar to him which probably caused a reduction in indexing time over what would have been expected through additional indexing experience.

By comparing the indexing performed by the same indexer of the same document, a statistical correlation coefficient can be obtained for essential terms which serves as a measure of intraindexer reliability or consistency. Likewise the common essential terms can be obtained between two experienced indexers as a measure of interindexer consistency.

The statistical correlation techniques used are the phi coefficient (ϕ) and chi square (X^2). The phi coefficient is particularly suitable for determining the degree of association between two dichotomous variables, that is,

variables which can be divided into two portions. The value of the ϕ coefficient is an indication of the degree to which the two dichotomous variables are associated. The ϕ coefficient can also be applied to continuous variables which are arbitrarily dichotomized.[4] For example, one may wish to determine the correlation, if any, which exists between family income and ownership of common stocks. The ownership of common stocks is already dichotomous since a family either owns stock or does not own stock. The family income can be dichotomized by arbitrarily setting a level of $10,000 and considering family income greater than $10,000 and less than $10,000. There are four possible situations which could occur: 1) families with income greater than $10,000 who own stock; 2) families with income less than $10,000 who own stock; 3) families with income greater than $10,000 who do not own stock; 4) families with income less than $10,000 who do not own stock. If one were to survey 100 families regarding income and stock ownership, the families could be categorized into the four categories. A 2 x 2 table could be constructed as follows:

INCOME

		< $10,000	> $10,000	
STOCK OWNERSHIP	Own Stock	(A) 15	(B) 40	55
	Do not Own Stock	(C) 35	(D) 10	45
		50	50	100

This table shows that of the 50 families of income greater than $10,000, 40 own stock, 10 do not own stock. Of the families of income less than $10,000, 15 own stock, 35 do not own stock. The table indicates that there may be a correlation between family income and stock ownership. The ϕ coefficient is a quantitative expression of the degree of correlation. Labeling the cells of the paradigm A, B, C and D, the ϕ coefficient is calculated as follows:

$$\phi = \frac{BC - AD}{\sqrt{(A+B)\,(C+D)\,(A+C)\,(B+D)}}$$ (Reference 4)

$$\phi = \frac{(40)\,(35) \;-\; (15)\,(10)}{\sqrt{(55)\,(45)\,(50)\,(50)}} = \frac{1250}{(50)\,(49.7)} = 0.503$$

The maximum ϕ coefficient would be 1.0. This ϕ coefficient would be obtained if all families have income greater than $10,000 owned stock and none of the families with income less than $10,000 owned stock. The paradigm would appear as follows:

INCOME

STOCK OWNERSHIP		< \$10,000	> \$10,000	
	Own stock	(A) 0	(B) 50	50
	Do not own stock	(C) 50	(D) 0	50
		50	50	100

$$\phi = \frac{BC - AD}{\sqrt{(A+B)\ (C+D)\ (A+C)\ (B+D)}} = \frac{2500}{2500} = 1.0$$

Chi square is a test of statistical significance. A significance test is a mathematical procedure applied to empirical data for deciding, on grounds of probability, whether or not a suitably formulated hypothesis is tenable. An observational result is said to be significant, and the hypothesis in question is rejected, when the obtained result belongs to an objectively specified unfavorable class having a fixed, small probability of occurrence in random samples from the hypothesized population.[11] In the case of testing the stock ownership - family income correlation, the following hypothesis is made: stock ownership is not dependent on family income. The alternate hypothesis is: stock ownership is dependent on family income. The obtained result must belong to the objectively specified unfavorable class having a fixed small probability of occurrence (level of significance) in random samples from the hypothesized population. If such a result is obtained, the hypothesis is rejected and the alternative hypothesis is accepted, i.e. stock ownership is dependent on family income. Usually the probability at which the results are significant is indicated with the results.

Chi square can be used to determine "goodness of fit" of data. In the example given above, a ϕ coefficient of 0.503 was obtained. The maximum possible ϕ coefficient was 1.0. One might ask whether the observed results represent a statistically significant "goodness of fit" to the ϕ coefficient of 1.0. Chi square (X^2) is related to the ϕ coefficient by the equation $X^2 = N\phi^2$ where N is the total number of individual items, in this case the number of families.

Calculating X^2: $X^2 = N\phi^2$ (Reference 4)

$$X^2 = 100\ (.503)^2$$

$$X^2 = 25.3$$

Referring this value to a statistical table of X^2 it can be determined that this result is significant at a probability (p) > 0.999. Therefore there is statistical evidence to support the hypothesis that stock ownership is correlated with family income. Referring to Figure 1, it can be seen that the X^2 value of 25.3 for N=100 falls on the curve above the line indicating p=0.999.

The ϕ coefficient and chi square can be employed to determine the correlation between essential index terms. Chi square has been used previously to determine the degree of term association i.e. the extent to which terms appear together in indexing.[9] Once judgments have been made of essential terms contained in a list of terms, the remaining terms are necessarily nonessential terms. Therefore, the index terms constitute a dichotomous variable. However, prior to the decision of whether a given term is essential or not, the decision itself is a subjective matter dependent on the experienced indexer who must apply his judgment, taking into account the essentiality of the term in the context of the report and in relation to the particular retrieval system. In the example of income and stock ownership, the ownership of stock is not at all subjective nor is the amount of income.

The essential and nonessential terms are compared by groups and the degree of correlation is determined between groups. For intraindexer consistency, the first group of terms is taken from one year's indexing and compared with the second group of terms from the other year's indexing, all taken from the same set of documents. The results are given in Table I and in Figure 1. The calculations are presented in Appendix II. The results show statistically significant correlation for essential terms regarding intraindexer and interindexer consistency.

Referring to Figure 1, it can be seen that the values for X^2 lie well above the p=.999 level of significance. The value of X^2 which represents statistical significance at p=0.999 is 10.83.[4] The curves representing X^2 as a function of the ϕ coefficient show that the number of terms included in the study has a considerable effect on the relationship. The degree of correlation (ϕ coefficient) for a large number of terms can be relatively small and still be statistically significant. On the other hand, with a small number of terms, a fairly high degree of correlation is required for statistical significance. The curves also demonstrate that as the degree of correlation increases, i.e., as the ϕ coefficient increases, the corresponding X^2 increases at a much faster rate. Thus a small incremental improvement in the degree of correlation gives a large incremental increase in the level of significance. These results tend to confirm that an experienced indexer's indexing with judgments of essential terms can serve as a reference standard for the evaluation of indexer trainees. The comparisons made in the remainder of this section will be based on the indexing of FLS in 1966 as a reference.

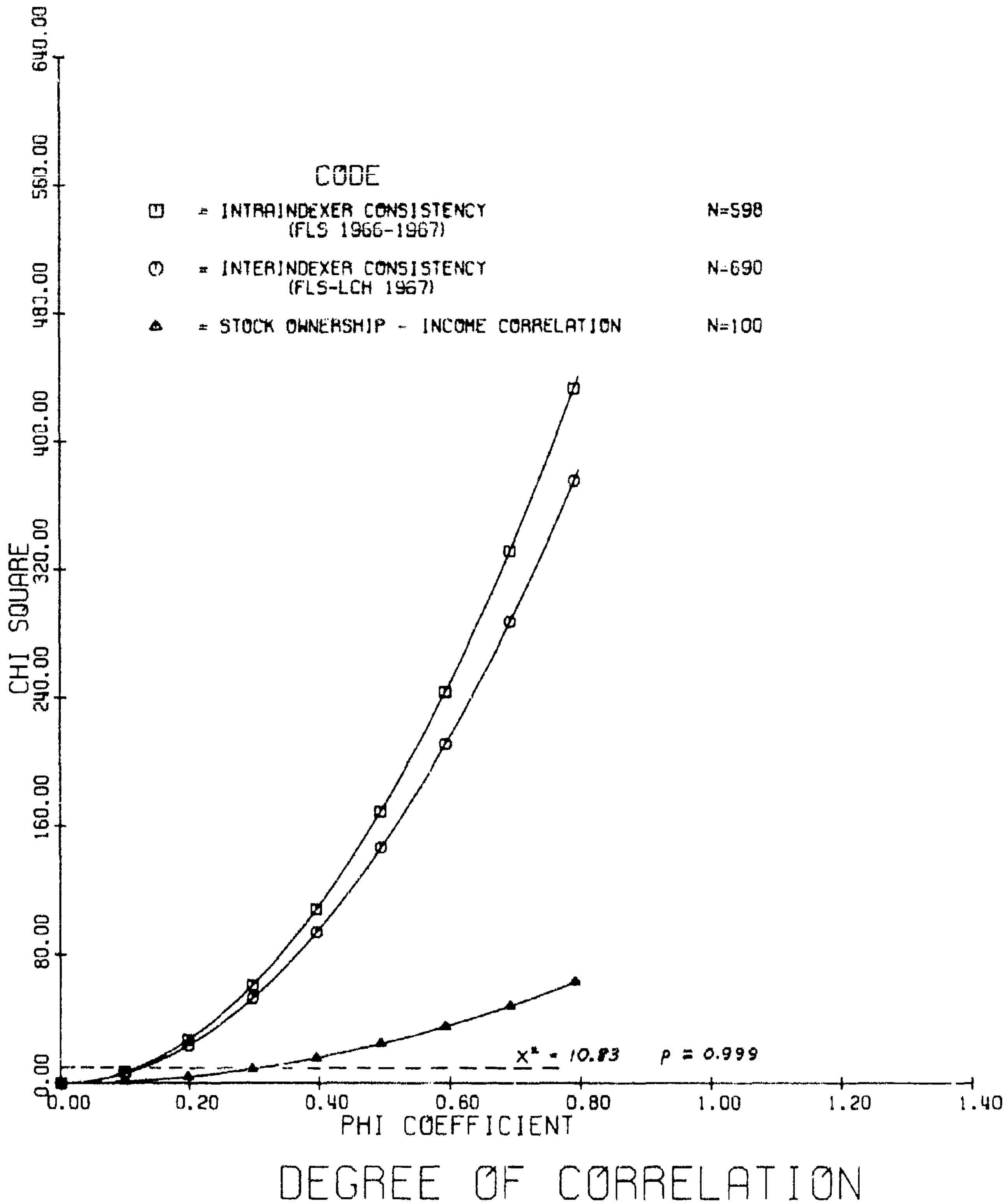

FIGURE 1. CHI SQUARE AS A FUNCTION OF PHI COEFFICIENT

TABLE I. INTRA - AND INTERINDEXER PHI COEFFICIENTS AND CHI SQUARE

Intraindexer FLS 1966 - 1967	ϕ Observed	ϕ max	Chi Square	Significance Level
Essential Terms	0.461	0.934	127	$p > .999$
Interindexer LCH 1967 - FLS 1967				
Essential Terms	0.327	0.814	73.8	$p > .999$

An interesting observation was made regarding intraindexer consistency. With consideration of terms judged essential in 1966 and terms judged essential in 1967, 73% of the terms judged essential in 1967 matched the essential terms of 1966. However, considering the entire list of terms from 1967 without regard to their judged essentiality, 86% of the essential terms of 1966 were included in the indexing of 1967. If the procedure is inverted, i.e., 1967 is taken as the reference standard, the corresponding figures are 68% and 92%. The results indicate that there is a "core" of essential terms (essential terms of 1966 plus essential terms of 1967, minus duplicated terms), which was used in indexing both years, although there is some discrepancy regarding the judgment of essentiality of terms within the core between the two years. The situation is illustrated in Table II. It should be noted that intraindexer consistency would be expected intuitively to be greater than interindexer consistency.

Figure 2 and Table III show the comparisons of all indexers involved in the study from both years for the period immediately following the training period. The improvement over the summer period made by the indexers of 1966 is indicated by dotted line extensions in Figure 2. The improvement was determined by performing an experiment with another set of 35 documents at the end of the summer. Appendix III presents the indexing of all indexers involved for a typical document.

Certain trends are evident in the data obtained. In the first place, the experienced indexer of both years used more terms for indexing in 1967 than in 1966. Slightly more terms (about 7% more) were indicated as being essential, although the percentage of terms considered essential based on the total index terms decreased from 48% to 43%. This finding further verifies the stability or reliability of essential terms and indicates that more nonessential or "helpful but not absolutely essential" terms were used in indexing in 1967. The time required for indexing was reduced considerably even though more index terms were used. It was indicated earlier that the indexer felt some familiarity with the material even though a year had passed between indexing and reindexing.

In comparing the performance of the two experienced indexers in 1967, it can be seen that one of them (LCH) used more index terms and considered more terms essential than the other (FLS). The sampling and comparison of individual index cards show that the terms judged essential by LCH are nearly identical to the "core" of essential terms derived from independent judgments of essential terms by FLS. This fact enhances even further the validity of essential terms as a reference standard. The time for indexing the documents was nearly identical for the two experienced indexers in 1967.

The graduate students were decidedly better in their performance than the students in 1966. Their indexing time was remarkably less and their index terms were somewhat more in agreement with the experienced

TABLE II. COMPARISON OF INDEXING OF DOCUMENT 29748 FOR THE SAME INDEXER IN 1966 AND 1967

29748 UNIFORM VACUUM ULTRAVIOLET REFLECTING COATINGS ON LARGE SURFACES (Ref. 5)

Indexer: FLS; May 1966	Indexer:FLS; July 1967	Core Terms
** Reflective coatings	**Reflective Coatings	**Reflective coatings
Vacuum	**Reflectance	**Reflectance
** Magnesium fluoride	**Ultraviolet radiation	**Ultraviolet radiation
** Aluminum	** Vapor deposition	**Vapor deposition
oo Coating	ooFilms	**Magnesium fluoride
** Mirrors	**Magnesium fluoride	**Aluminum
** Ultraviolet radiation	**Aluminum	**Mirrors
** Reflectance	**Mirrors	o*Surfaces
Measurement	Aluminum coatings	
** Vapor deposition	Wavelength	
o Surfaces	Protective coatings	
oo Films	ooCoating	
	Manufacturing technology	
	* Surfaces	

CODE:

** Judged essential both years

* Judged essential only in one year

o Essential term of one year included in indexing of other year as nonessential term

oo Common nonessential terms

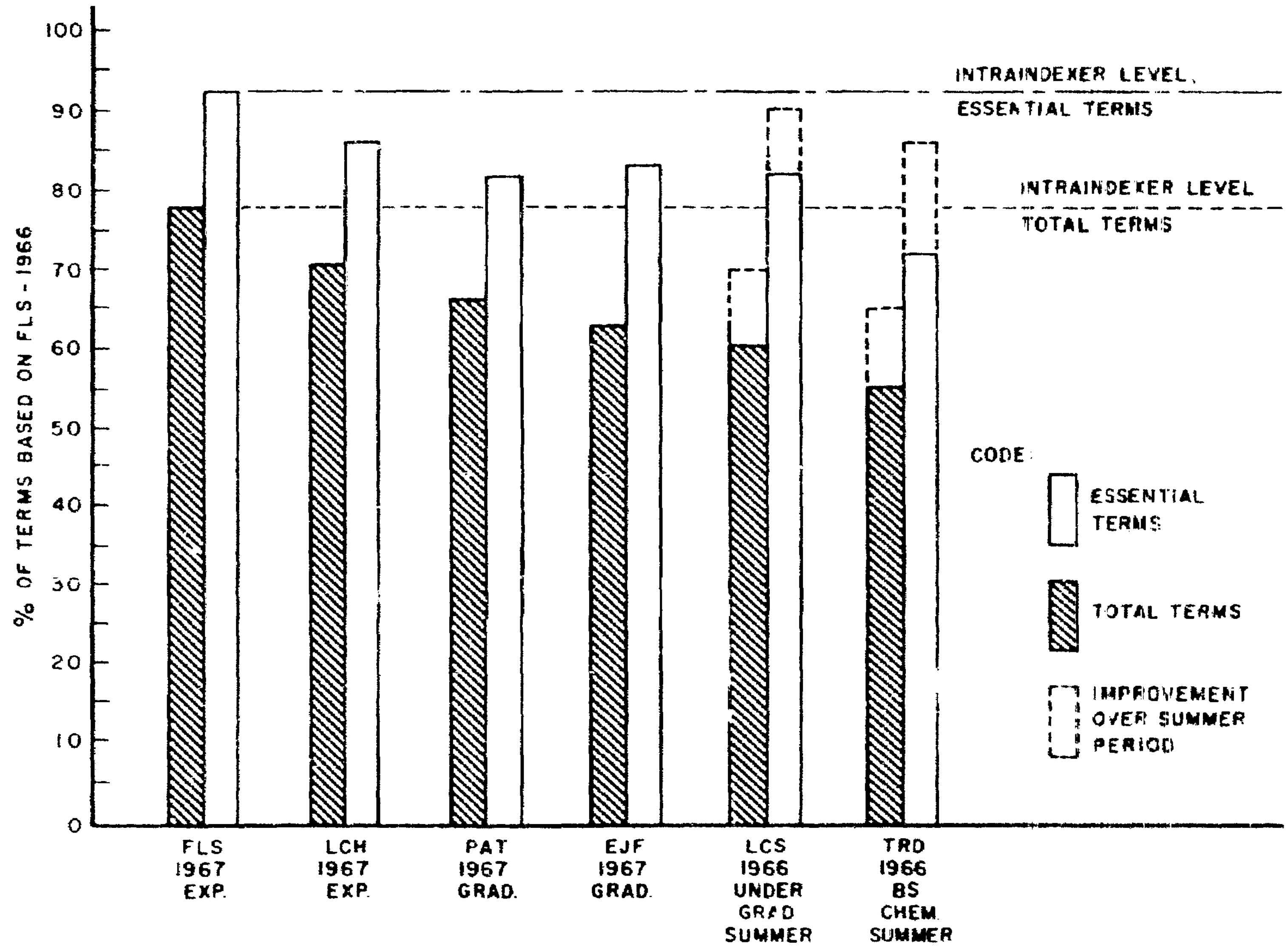

FIGURE 2: COMPARISON OF INDEXERS RELATIVE TO EXPERIENCED INDEXERS (FLS - 1966)

TABLE III. COMPARISON OF INDEXER TRAINEES WITH EXPERIENCED INDEXERS

Indexer	Ave. Total Terms Used in Indexing	Ave. No. of Terms Judged Essential	% Common Total Terms (FLS 1966)	% Essential Terms Included in Indexing (Only FLS Essential Terms considered	Ave. Misleading Terms Used in Indexing	Ave. Time for Indexing (Minutes)
FLS (66)	15.5	7.5	100%	100%	————	25.4
FLS (67)	18.6	8.0	77.5%	92.0%	————	14.3
LCH (67)	20.6	10.3	70.0%	85.2%	————	14.5
PAT (67)	22.0	——	65.7%	82.0%	0.34	24.0
EJF (67)	20.7	——	62.5%	82.6%	0.10	17.4
LCS (66)	15.4	——	60.0%	82.0%	0.63	8.6
TRD (66)	12.2	——	54.0%	71.1%	0.51	7.0

FLS Experienced indexer both years

LCH (67) Experienced indexer

PAT
EJF Graduate-Research Assistants

TRD BS Student - summer employee.

LCS (66) Undergraduate student - summer employee

indexer regarding both essential terms and total terms. These results can be explained largely on the availability of a better thesaurus, the more extensive background of the indexer trainees, and the modification of the training program. Other possible factors may also help to account for the improved performance as indicated in Section II.

It is of interest to consider why indexers, and especially, the experienced indexers, tended to use more total terms and essential terms in indexing in 1967 than in 1966. As stated earlier, the retrieval of a document abstract gives the requestor the opportunity for acceptance or rejection. Nonretrieval of an item does not afford the requestor the opportunity of seeing it to make such a decision. More index terms increase the probability of retrieval. Another requestor trend is the desire for specific information which necessitates indexing of specific concepts. Both of these requestor parameters have caused the philosophy of indexing to tend somewhat towards the use of more indexing terms and perhaps a somewhat broader interpretation of what essential terms are.

Even though it sounds paradoxical, less time is required in indexing with more terms. This phenomenon is clearly indicated in the results. Probably the explanation is that decision making regarding terms is reduced. One need not be quite as selective in the index terms used. If there is some doubt regarding the use of a particular term, the term is assumed to be of sufficient merit to be listed. With more selective indexing, the decision of whether to list it may be fairly time consuming, and these decision making time increments are additive.

Another reason for the improved performance is that the thesaurus was revised between the summer of 1966 and the summer of 1967. The revised thesaurus contains a much larger number of reference terms which indicate how to express concepts in active vocabulary terms. The improved thesaurus undoubtedly was beneficial to the indexer trainees.

SECTION IV

DOCUMENT RETRIEVAL SYSTEM OPERATION AND AMIC RELATED ACTIVITIES

Input. One of the primary objectives of the Information Systems Section concerning the AFML contract is the indexing of newly acquired documents for inclusion into the document retrieval system and the retrospective searching of the file in response to search requests. Two changes made in the input operation have improved efficiency. Formerly, keypunch cards, which are used for updating the computer search tape, were processed by the keypunch section of the Research Institute. During the summer the availability of keypunch equipment suggested the use of our own clerical personnel for keypunching. Keypunching by Information Systems Section personnel has proved effective. Familiarity of the clerical personnel with the technical vocabulary and interest in their work has kept keypunch errors to a minimum. Further, tighter control on the status of keypunching is maintained. Another change which has proved beneficial is the inclusion of the Defense Documentation Center (DDC) AD number or the NASA N number with each abstract. This information is cross-referenced at the AFML Library, but it is more convenient to have the AD or N number available immediately on retrieval of the document abstract. The document can then be ordered if desired from DDC or the NASA information facility.

Microfilm (16 mm) reels are being prepared of all abstract and index cards. As indicated in Reference 10, the use of microfilm provides a permanent accessible record and facilitates the screening of searches. As microfilming has proceeded, another advantage has become evident. The preparation of microfilm requires accurate and complete records of index cards and abstracts. Microfilming has revealed a number of cases of incompleteness which have been brought up-to-date. Microfilming thus serves as an input control function as well as providing a permanent record.

During the period covered by this report, 1 December 1966 through 30 November 1967, approximately 4250 documents were indexed and processed into the system. The documents were indexed with an average of 24 terms per document (exclusive of automatic generic postings) with an average index time of about 25 minutes. Distribution by subject category is shown in Appendix V. There are now approximately 31,000 documents in the AMIC document retrieval system.

Searching. A total of 286 technical requests were processed by the Information Systems Section last year. An average of 9.0 abstracts was printed per search from the microfilm records for inclusion as part of the search results. User feedback is very favorable toward the receipt of abstracts. A complete listing of searches processed by title, number,

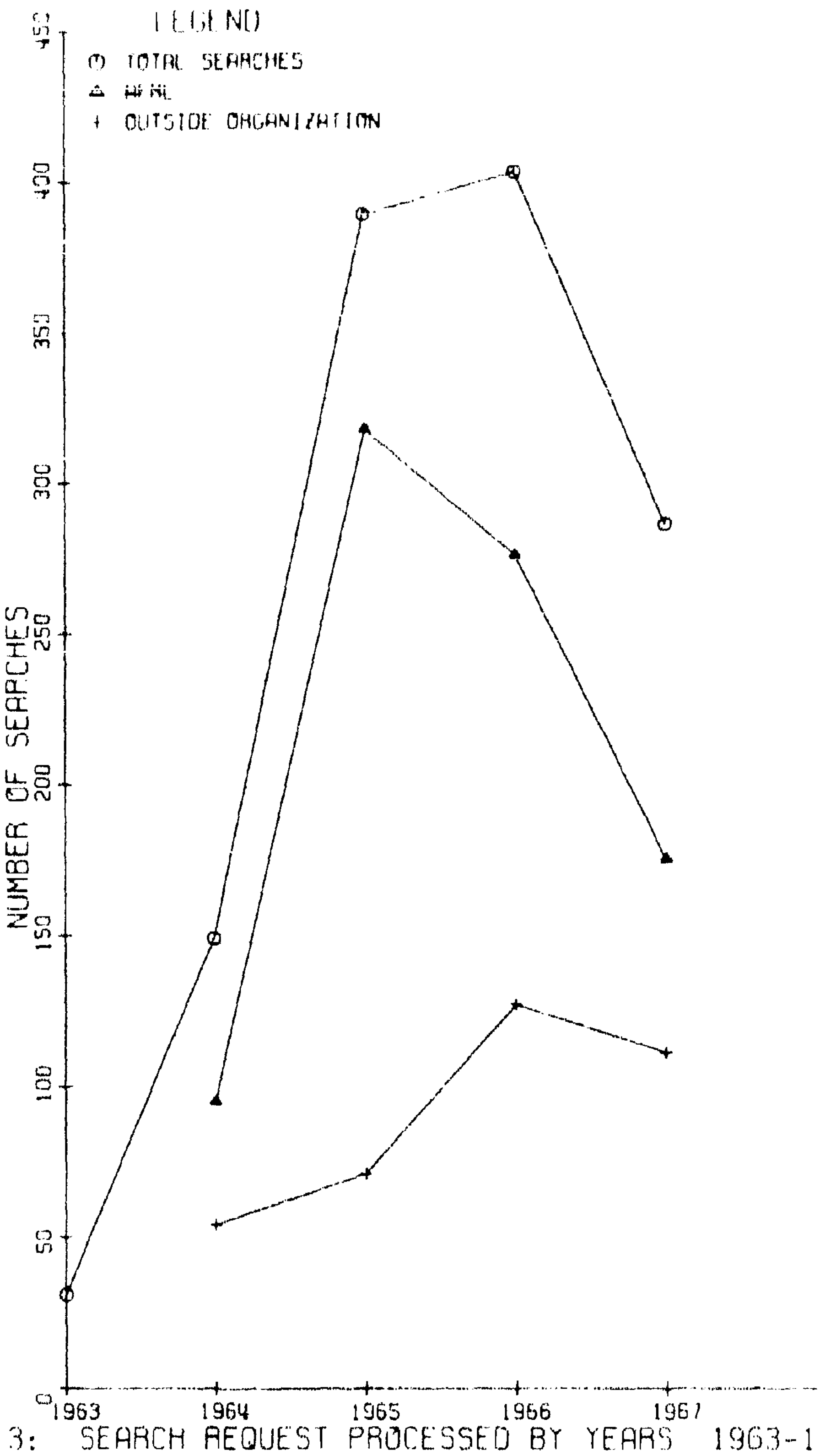

FIGURE 3: SEARCH REQUEST PROCESSED BY YEARS 1963-1967

requestor and organization is presented in Appendix IV. It includes an index of requestors and corresponding search numbers, and the number of searches requested by various organizations.

The number of technical requests received declined about 26.5% from last year. Figure 3 presents the total number of requests processed by the AMIC document retrieval system since 1963. The number of search requests performed for the AFML and for all other organizations is indicated in addition to the search totals. It can be seen readily from Figure 3 that the number of search requests from the AFML has been on the decline for the past two years. Whereas the number of search requests from other organizations has increased with only a slight decline this year. The drop-off in the number of AFML requests is a matter of some concern, since AFML personnel should be prime users of the AMIC system.

Several reasons can be advanced for the reduced usage of the AMIC system by the AFML. Every new project started in the AFML is required to go through AMIC before a request for proposal (RFP) is submitted for bids. The project engineer thus gets a background of the state-of-the-art of the topic he is pursuing. Such a search request may reveal that certain information is already available and that the RFP should be modified to prevent exact duplication of work already performed. On the other hand, the search may indicate by lack of retrievals, that research in the proposed area has not been done previously, thus underscoring the need for proceeding with the project. The number of new projects initiated in 1965 and 1966 was fortuitously quite high; in fact, such a level of new projects could not be expected to be maintained each year. Coupled with this factor is the reduced funding available for Air Force research due to the Vietnam conflict.

Another possible reason for the reduced number of requests for retrospective searches is that a more effective SDI program is operating. In addition to the Materials Information Bulletin which provides a Keyword in Context (KWIC) title list of new AFML reports, an SDI program has been initiated with Chemical Abstracts Service (CAS), and the AFML is currently participating in the NASA SDI program. The decrease of AFML search requests is an area that should be investigated.

A listing of search requests by subject category is presented in conjunction with a listing of documents input into the system by subject category. This parallel listing indicates that document acquisition is reasonably reflective of the desires of the users as indicated by search requests.

Miracode system. The AFML maintains a Miracode system at the Materials Information Branch. This system operates with optically coded 16 mm cartridge microfilm. Boolean logic operators can be applied in searching. Several auxiliary systems are maintained on the Miracode

system. The University provides assistance by keypunching items for eventual coding and inclusion in the appropriate system. The Commerce Business Daily (CBD) and independent research reports are two systems the laboratory maintains. Additionally, the computerized AMIC document retrieval system is duplicated on Miracode so that searches arriving at the Materials Information Branch requiring immediate response can be searched on site. The search capability, however, is not as sophisticated as is available with computer searching.

It is expected that handbooks, which are not now included as part of the AMIC system, will be assigned special access numbers, indexed, and included as a special Miracode file. Symposia will be kept at the AFML Library and retrieval provided by a KWIC index of the individual papers contained in the symposia proceedings.

Thesaurus. Several changes were made in the last updating of the thesaurus. To make the thesaurus more consistent with the practice of most thesauri, the PO, GT, RT, SA designations were converted to Broader Term, (BT), Narrower Term (NT), Related Term (RT), and also see (AS), respectively. The new designations have the additional advantage that sorting of these items alphabetically places them in the desired order.

A very useful entry introduced into the thesaurus is the "collection" term which serves to group similar items under a common heading. These are terms which are not hierarchically related, since the BT and NT designations effectively "collect" the hierarchically related items. An example of such a collection term is ADDITIVES. This term shows that "additives" is not an active vocabulary term but the active terms which could be considered additives are displayed in one group. The entry appears in the thesaurus as follows:

ADDITIVES - AS INHIBITORS INCLUDES -

CERAMIC STABILIZERS	0837550
CHEMICAL STABILIZERS	0878000
FUEL ADDITIVES	1554000
LUBRICANT ADDITIVES	2169500
PIGMENTS	3171500
PLASTICIZERS	3193000
POLYMER FILLERS	3248100
PRESERVATIVES	3314500

One problem which has tended to limit vocabulary and thesaurus changes, particularly additions, is the retrospective updating of previous documents. It is essential for any vocabulary to be amenable to modification to accommodate changes in technology and user requirements. The AMIC vocabulary has already undergone considerable modification as reported earlier.[6,7] However, as the vocabulary has become more stable and the store of documents has increased, there has been increasing reluctance to introduce changes. The retrospective updating of previous document indexing is at best difficult and at worst impossible. The dilemma can be illustrated by the following question: if a new active vocabulary term is introduced into the system, how can all previous documents which should have been indexed with the term be recalled and the new term added to the indexing? With some terms there is no satisfactory answer. On the other hand, the decision to make a necessary change (i.e., as shown by user requirements and/or technological evolution) should not be withheld simply because the new term cannot be unilaterally applied to all previous documents.

To cope with the problem, it was decided to make necessary changes (additions) and update previous documents whenever possible. If this is not possible, the date at which the term becomes effective is indicated in the thesaurus. This effective date informs the search strategist that to retrieve documents previous to the activation of a term, allowance must be made in the search strategy for this contingency. For example, MICROMECHANICS became effective 1 July 1967. To retrieve previous documents dealing with the concept, the search strategist must think of alternative terminology which might have been used to express it. Terms such as INTERACTIONS and REINFORCEMENT (matrix-reinforcement interactions to express micromechanics), MICROSTRUCTURE, PLASTIC DEFORMATION, and RATE EFFECT might be considered in the search strategy.

American Society for Metals Search Service. The University of Dayton had been participating in the Research Associate Program (RAP) of the American Society for Metals (ASM). The objective of the program was to submit searches to ASM for running on their computerized information retrieval system of metallurgical journal literature. The preponderant interest of AFML requestors in metallurgy and metallurgy related topics motivated the University's participation in this program to be able to serve those needs more effectively. Despite a number of ASM RAP participants, the service proved too expensive, and satisfactory arrangements for continued financial support were not forthcoming. The service has been forced to be discontinued.

Chemical Abstracts Service SDI Program. To provide improved information service to technical and professional personnel in the Air Force Materials Laboratory, a joint effort involving the Materials Information Branch (MAAM), the Polymer Branch (MANP), and the Information Systems

Section of the University of Dayton is under way in which the University is participating in CAS SDI programs for the benefit of selected AFML personnel. Three types of document abstracts are included: SDI Experiment, Polymer Science and Technology from journals (POST-J), and Polymer Science and Technology from patents (POST-P). A description of these SDI programs is given in Appendix VI in conjunction with typical user profiles.

Informal comments from the users are generally favorable regarding the CAS SDI program. The initial success achieved and the acquisition of more experience may lead to an expanded endeavor.

It is intended by CAS that the SDI programs and computer tapes be available to those users desiring to run their own user profile searches. The alternative is for CAS to run the programs on a service basis with a service charge dependent on computer time required. The programs and tapes should be capable of being run on the IBM 360 computer Model 40 acquired with the advent of a cooperative venture between McCall Corporation and the University. The SDI program will be run at the University to determine the cost comparison between running the programs and the service charge incurred by having them run at CAS. If it is shown that it would be more economical to run them at the University, this procedure would be continued.

Personnel time distribution. The time spent by various personnel has been broken down into various task numbers. The task number definitions are given in Figure 4 and the personnel time distribution for various types of personnel are presented in Table 4.

FIGURE 4

DEFINITION OF TASK NUMBERS

01	General	Includes:
		Supervision
		Meetings & trips
		Holidays & sick leave
		Writing of reports
		Training of students
		Time spent with visitors
02	Input	Includes:
		Assignment of accession numbers
		Document accounting records
		Preparation of index and abstract cards
		Indexing
		Keypunching
03	Output	Includes:
		Preparation of search strategy
		Search
		Screening of searches
		Search accounting records
		Library loan functions
		Keypunching
04	Updating	Includes:
		Review of vocabulary and thesaurus
		Changes or additions to previous records
		Keypunching
		Acquisition of missing documents
05	(UD) Research	Includes:
		Evaluation studies
		Studies of new techniques
		Investigation of new systems

FIGURE 4 (Continued)

05	(AFML) Library	Includes: Preparation of Materials Information Bulletin
06	Special Projects	Includes: Work performed in support of the AFML not directly related to AMIC retrieval system
07	Microfilming	Includes: Time spent on the microfilming of index/abstract records

TABLE 4

DISTRIBUTION OF PERSONNEL TIME BY TASK NUMBER

Professional and Clerical at UD:

Task Number	Percent of Time
01	23.4
02	46.4
03	11.8
04	4.3
05	1.8
06	10.5
07	0.8

Clerical at the AF Materials Laboratory Library

Task Number	Percent of Time
01	12.3
02	49.0
03	7.9
04	3.2
05	15.9
06	11.6
07	

SECTION V

ESTABLISHMENT OF AN INFORMATION RETRIEVAL SYSTEM FOR INFORMATION SCIENCE RELATED DOCUMENTS

The AFML Library has received a number of information science documents. These documents had not been indexed previously because all available indexers were required for indexing materials documents for the AMIC document retrieval system. It was agreed between the AFML and the University that the inclusion of information science documents in the AMIC document retrieval system would be mutually beneficial.

In the summer of 1967, the Department of Information Science was formally inaugurated at the University of Dayton. The Information Systems Section has employed a graduate student from the Information Science Department to index the information science related documents. The training of this indexer has not been nearly as readily accomplished as the training of students for indexing materials documents (see Section II) since no standard information terminology nor thesaurus has been derived for information terminology. Training has consisted primarily of exposure to literature on coordinate indexing and attempts at indexing with a review.

Presently "free indexing", i.e., selection of "index terms" without reference to a predetermined vocabulary or thesaurus, is being performed using the entire document as the source for indexing. This free indexing is being done by the information science student and being reviewed by an experienced indexer (experienced, however, primarily in the area of materials documents). The authors' terminology is retained to a large extent unless there are good reasons to modify or supplement it. A group of several hundred documents will be free indexed subsequent to which the "terms" will be keypunched, sorted, and critically reviewed. This review will be directed toward the grouping of similar concepts with relationships of terms (generic or hierarchical relationships as well as "related term" concepts) and to the selection of actual vocabulary terms which most appropriately define and express the concept involved. This procedure was the one followed at the beginning of the University's involvement with the AFML in the establishment of the AMIC document retrieval system[5] and the technique was also applied in a study for the Air Force Flight Dynamics Laboratory.[3] It is anticipated that the terminology of information science may present problems since information science is a newly recognized discipline and is generally more abstract than the area of materials. However, the experience acquired from previous efforts in vocabulary establishment and thesaurus development will be invaluable in deriving a useful vocabulary and thesaurus from the free indexing which will serve as the basis for inclusion of information science documents in the AMIC document retrieval system.

SECTION VI

SUMMARY

Two graduate students in chemistry and a high school physics teacher were employed for the indexing of technical documents. Their training followed closely the indexer training program devised the previous year for two chemistry students. Several modifications were made which improved training procedures. It was found that de-emphasis of the extracting function, i. e., writing out of concepts in word, phrase, or sentence form, was helpful in reducing indexer training time and in permitting consideration of the document as a whole. The document analysis technique was found to be very important in learning the indexing function. Modification of the training program resulted in a reduction of the formal training period from four weeks to three weeks, and the results obtained were better than had been achieved the preceding year, particularly regarding indexing time.

A study of intra- and interindexer consistency was conducted. Chi square and the phi coefficient (see Appendix II) were the statistical techniques employed to determine the correlation of indexing of both total and essential index terms. It was discovered that the intra- and interindexer correlation for experienced indexers was statistically significant at $p > 0.999$. This high degree of consistency tended to verify that an experienced indexer's indexing can serve as a reference standard for evaluating the work of indexer trainees. The agreement of essential terms was markedly better than was obtained for total terms. Using one indexer's set of index terms as a reference point, comparisons were made of the performance of all indexers involved in the study. A graphical representation (see Figure 2) shows that with experience a high degree of consistency among indexers can be expected under the conditions provided, i.e., materials type documents as an indexing source and an established thesaurus.

In the operations of the document retrieval system, it was noted that the number of technical search requests declined over the past year. A graphical representation of the search requests (see Figure 3) submitted by both the AFML and all other organizations indicates that the drop-off is attributable primarily to reduction of use of the system by AFML personnel. Since the procedure for initiating new projects results in search requests for the document retrieval system, the usually high number of new AFML projects in 1965 and 1966, as compared to 1967, may account in large measure for the decline in AFML search requests. More effective SDI programs may also be affecting the level of use of the AMIC document retrieval system. Minor thesaurus changes were effected. The ASM retrieval service was forced to be discontinued, and thus is no longer available to AFML personnel.

An SDI program with Chemical Abstracts Service was initiated for the benefit of AFML personnel. SDI Experiment, POST-J, and POST-P are the CAS SDI services subscribed to. Preliminary results are encouraging, but profile modifications are being made to improve the service.

The Materials Information Branch has a number of information science documents which are to be included in the AMIC document retrieval system. A graduate student in Information Science has been employed for indexing these documents. Free indexing is presently being performed from which an active index vocabulary and thesaurus can be derived which will lead to inclusion of information science documents into the system. Assistance in establishing an information science system may be forthcoming from anticipated contacts with other organizations involved in this area.

SECTION VII

FUTURE WORK

A new concept in the indexing of technical documents with computer assistance has been evolved in principle and will serve as an area of research for the coming year. The ultimate objective of the research is the development of a computer-based information retrieval system which can be accessed directly by the user, i. e., by the requestor of information without the requirement of specialized training in the use of the system.

In essence, the procedure is to enter the entire text of abstracts (other portions such as the table of contents, conclusions, selected specific terms, etc., will be studied later) into the computer. The computer will order the words alphabetically and compare each word with two word lists. One word list will contain nonsignificant words, e.g., the, and, or, studies, etc.; the other word list will contain technical words and the appropriate disposition of those words. This technical word list may be considered to be a computer contained thesaurus which will channel all technically meaningful words to the appropriate active vocabulary terms and the correct generic postings. Words of technical nonsignificance from the abstract which match the nonsignificant word list will be disregarded. Words which do not match either list will be stored and printed on demand for decisions. The information specialist would then have the assignment of deciding if the nonmatching terms should be entered into the nonsignificant word list or the technical term list, and, if it is to be entered into the latter, how should it be entered, what generic relationships exist, what related terms apply, etc. The scheme may be considered as machine aided indexing in which the processing capability of the computer is integrated with experienced human intellect. The repetitive operations which are required in human indexing would be assumed by the computer, while the important problems of term relations, semantics, etc., would be concentrated on by the information specialist.

On the output side of the operation, a request stated in normal text English would be entered into the computer and processed in the same way as the text abstract, i. e., with nonsignificant terms disregarded as instructed by the nonsignificant term list and the technical terms channeled to the appropriate active term locations. A means of applying Boolean logic to the terms will be devised to coordinate terms in an appropriate sequence. Interaction between the requestor and the computer would take place so searching and retrieval can be optimized.

To limit the scope of the endeavor, it was decided to concentrate on the area of composite materials and to consider a population of about 500 documents, large enough to be representative, but small enough to be amenable to

treatment. One of the objectives to be met by the program is an evaluation of the machine aided indexing and retrieval system by comparing it to the present system wholly derived from human indexing. Another objective is to indicate the feasibility of whole text processing both for input and for retrieval with allowance for direct requestor-system interaction.

To carry out the research program outlined above, various hardware and software capabilities must be established. Considerable software must be prepared before the hardware is needed. It is expected that a remote Selectric typewriter terminal to the IBM 360, Model 40, will be available in the offices of the Information Systems Section by January 1968. This rather rudimentary equipment communicates with the computer character by character, but this equipment should prove adequate for the early phases of the study. By the spring of 1968 it is hoped to have a cathode ray tube (CRT) terminal available. This unit will provide much greater flexibility in display and message transmission to the computer; up to 480 characters can be accumulated, corrected, and transmitted at once to the computer. By the summer of 1968, it is possible that a remote typewriter unit could be available at the Air Force Materials Laboratory. The actual installation of such a unit at WPAFB would be dependent on the progress made at UD in developing the system which in turn will determine the usability of the system by users at WPAFB.

REFERENCES

1. Anonymous, Preparation of Search Profiles, Chemical Abstracts Service, Columbus, Ohio, Februa 1967.

2. Anonymous, SDI Experime Selective Dissemination of Information, Chemical Abstracts Servic Columbus, Ohio, 1967.

3. A. Debons, F. L. Scheffler, J. D Snide, Development and Experimental Evaluation of a Retrieval System for Air Force Control-Display Information, AFFDL-TR-67-119, Air Force Flight Dynamics Laboratory, WPAFB, Ohio, July 1967.

4. G. A. Ferguson, Statistical Analysis in Psychology and Education, New York: McGraw-Hill Book Company, Inc., 1959.

5. H. Herzig, Uniform Vacuum Ultraviolet Reflecting Coatings on Large Surfaces, NASA TN D-3357, National Aeronautics and Space Administration, March 1966.

6. E. A. Janning, Establishment of a Coordinate Indexing Retrieval System for the Air Force Materials Laboratory, RTD-TDR-63-4263, AD 428423, Air Force Materials Laboratory, WPAFB, Ohio, November 1963.

7. E. A. Janning, The Modification of an Information Retrieval System by Improving Vocabulary Control, Indexing Consistency, and Search Capabilities, AFML-TR-65-20, AD 613301, Air Force Materials Laboratory, WPAFB, Ohio, March 1965.

8. E. A. Janning, Operations of a Document Retrieval System Using a Controlled Vocabulary, AFML-TR-66-36, AD 633614, Air Force Materials Laboratory, WPAFB, Ohio, March 1966.

9. P. E. Jones and R. M. Curtice, "A Framework for Comparing Term Association Measures," American Documentation, Vol. 18, No. 3, (July 1967) pp. 153-161.

10. F. L. Scheffler, Student Indexer Training Program and the Improved Operation of a Document Retrieval System, AFML-TR-66-391, AD 651039, Air Force Materials Laboratory, WPAFB, Ohio, January 1967.

11. G. P. Wadsworth, Introduction to Probability and Random Variables, New York: McGraw-Hill Book Company, Inc., 1960.

APPENDIX I

BACKGROUND INFORMATION FROM PREVIOUS REPORTS

The original system established by the University was a coordinate indexing system based on the use of very specific terminology and the utilization of links and roles. The information to be contained in the system was a collection of approximately 10,000 scientific and technical reports that dealt with all aspects of materials research of interest to the Air Force. There was no predetermined vocabulary and the indexers were permitted to use whatever terminology that was needed to describe adequately the documents they indexed. The first five hundred documents that were indexed averaged over eighty link-role-term combinations per report. As the indexers gained experience, the average number of entries was gradually reduced to slightly over forty.

An attempt to build a thesaurus of the system's vocabulary was made during the second year of operation. At that time, about 6,000 documents had been indexed and a thesaurus of slightly over 18,000 terms was generated When the system was updated in December 1963, the addition of 4,000 documents to the system generated 30,000 new terms. Although the system had been used to effectively retrieve information in answer to search queries, the rapid growth of the vocabulary became of concern since it was desirable at that time to maintain manual searching capabilities.

The results of an evaluation of the use of links and roles in the system, and the rapid growth of the vocabulary, formulated the basis for modifying the system during the fourth year of operation. Although the use and value of role indicators were questioned during the period of establishing the system, a decision was made to continue their use until the system was operable and could be formally evaluated. An evaluation was made after the system became operable and confirmed the opinion that the use of roles was more detrimental than helpful to the system. The use of roles was then discontinued and that elimination, in addition to the continual growth of the vocabulary, necessitated modifications in, and the establishment of, vocabulary controls.

The naming of materials, primarily metallic alloys and organic compounds, accounted for the major portion of the rapid and apparently unending growth in the vocabulary. Since the documents that were indexed into the system dealt with the research and development of materials, the generation of new terms was anticipated but the quantity far exceeded the expectations. The approach taken to establish a vocabulary control was to divide the vocabulary into classes of materials for analyses. The purpose of each analysis was to determine the possibility of generalizing the naming of materials without losing an undue amount of specificity. Through these generalizations, controls

were established for the naming of materials and the system's vocabulary was reduced to approximately 10,000 terms. Although 6,000 new documents had been added to the system during 1965, the vocabulary remained fairly constant. The use of the controlled vocabulary was found to be very beneficial to the overall operation of the system.

The continued growth and increased searching of the system necessitated additional indexing personnel. Efforts to fill these positions with professional persons were unsuccessful. It was then decided to employ students with technical backgrounds to index technical documents.

Two students having backgrounds in chemistry and one physics major were employed and after completing a training program, they became proficient indexers. The training program was specifically designed to include all aspects of information retrieval, although the main emphasis was on the indexing function. By being exposed to the overall system, the students acquired an appreciation of the importance of good indexing. The students were supervised by professionals well versed in indexing and the other aspects of information retrieval and having similar technical backgrounds. The training program was flexible to allow the students to proceed without excessive pressure. The students were expected to demonstrate initiative in their own advancement as indexers. About four weeks of full time work were required to complete the training. The training program consisted of the following steps:

1. Introduction to information systems
2. Document analysis
3. Extracting
4. Thesaurus use
5. Exposure to screening of searches
6. Direct indexing
7. Evaluation

The training program was well accepted by the students as evidenced by their written comments. They felt they had benefitted by their experience. It was concluded that students with a technical background are capable of becoming proficient indexers by applying the training program outlined above.

Several thesaurus changes were effected in the thesaurus updating (November 1966). "SEE" terms were changed to "USE" terms. A number of reference terms were entered which show which active vocabulary term(s) are used to express a concept. An example is: HIGH ENERGY RATE FORMING USE FORMING 153800 and RATE EFFECT 3408900. Scope note type terms were added. These entries explain the scope encompassed by an active term or define a term if a definition is needed.

Dates of issue of documents were included as active index terms so that searching by date is possible. Subject categories were established based on Committee on Scientific and Technical Information (COSATI) categories to provide data on the subject distribution of new acquisitions. This will be useful in acquisition control and will give a quantitative measure of the contents of the system. Microfilm printer-readers and microfiche readers were installed at the University of Dayton. The use of this equipment has resulted in better records and greater efficiency in operation.

APPENDIX II

CALCULATION OF PHI COEFFICIENT AND CHI SQUARE FOR INTRA- AND INTERINDEXER CONSISTENCY

<u>Calculation of the degree of association of judgment of essential terms</u>

It was stated in Reference 10 that the degree of correlation between the students' indexing and the supervisor's indexing was better with regard to essential terms as compared to the correlation of total terms. It was inferred from this result that essential terms as judged by an experienced indexer represent a better reference standard for indexing evaluation than the total index terms used. Since the same indexer indexed the same documents one year later, a measure of the reliability of the indexer's judgment could be obtained. The criterion for establishing the reliability of the experienced indexer's judgment was the degree of association or correlation between the judgments of essential terms made independently one year apart. If it can be shown that there is a statistically significant association, then one has evidence for the verification of the hypothesis that an experienced indexer's judgment of essential terms is reliable. The particular statistical technique chosen was the determination of the ϕ coefficient and chi square.[4]

One factor which was considered was the possibility that the judgment of essential terms could be misleading if the basis of the study were chosen so that random operants might cause a fortuitous preassociation. Such an occurrence would detract from the interpretation of the results unless random preassociation were accounted for. For example, suppose a document were indexed with five terms, four of which were judged to be essential by an experienced indexer. If the document were given to a second experienced indexer, and the second experienced indexer selected the same four terms as essential terms, the ϕ coefficient would be 1.0, indicating perfect correlation. But by chance alone, the probability of the first term being an essential term is 4/5. The probability of the second term being an essential term is 3/4. The probability of the first two terms being essential is $4/5 \cdot 3/4 = 0.60$. The probability of all four terms being randomly selected as essential terms is $4/5 \cdot 3/4 \cdot 2/3 \cdot 1/2 = 0.20$. Therefore, there is some chance that random preassociation may occur. The observed data indicated that the average document indexing consisted of about 20 terms of which 9 were judged essential. The probability that all 9 essential terms would be selected randomly from the list is calculated as follows: $9/20 \cdot 8/19 \text{------} 1/12 \cong .0006$. This value is so close to 0 that for all practical purposes it may be stated that random preassociation does not occur.

Observed degree of association of essential terms - intraindexer consistency

To test the reliability of judgment of essential terms, the same list of terms should be provided to several judges (or, in this case, the same judge one year apart), and the degree of association determined. It was observed that about 45% of the terms were judged to be essential by FLS both in 1966 and 1967. If the judgments were perfect, for a typical document of 20 terms, 9 terms would have been judged essential in 1966 and the same 9 terms would have been judged essential in 1967. It was shown above that there is no expected significant association of terms if 9 terms are randomly assigned out of a possible 20 terms for two different judgments. In actual fact, the lists of terms to be judged were not identical. The number of terms judged essential was about 7% higher in 1967 than in 1966. The total number of terms used and hence, forming the basis for judgment, was about 20% higher in 1967 than in 1966. If one can assume, however, that the total number of terms used in indexing does not appreciably affect the judgment of whether or not a term is essential, an assumption already inherent in the concept of essential terms, one can determine the average number of total terms used for indexing the series of 35 documents in the two years. This figure can be used for calculating the ϕ coefficient and X^2 for obtaining the degree of association of the essential terms over the two years. It is necessary that the same number of total terms be used in the calculations. It should be noted that the actual essential terms obtained were used. The total terms for 1966 and 1967 were adjusted to the average for calculation purposes. Since the number of essential terms of the two years is not equal, the maximum possible ϕ coefficient is less than 1.0. The ϕ coefficient obtained for the association of the essential terms is a good indication of the reliability of the judgment of essential terms.

Observed data

		1967 Non. Ess.	1967 Ess.	
1966	Ess.	(A) 70	(B) 191	261
	Non Ess.	(C) 247	(D) 90	337
		317	281	598

$$\phi = \frac{BC - AD}{\sqrt{(A+B)(C+D)(A+C)(B+D)}} = \frac{47177 - 6300}{\sqrt{(261)(337)(317)(281)}}$$

$$\phi \approx 0.461$$

$$\chi^2 = N\phi^2$$

$$= 598\,(.461)^2$$

$$= 127$$

This value of χ^2 is significant at $p > .999$. Therefore, a high degree of association is indicated.

Maximum possible degree of association of essential terms

The maximum possible degree of correlation can be determined by considering the hypothetical case of all terms considered essential in 1966 also being considered essential in 1967. Referring to the table derived from the observed data, one can see that Cell B must now be 261 since all essential terms from 1966 are considered essential also in 1967. Cell A must be 0 since there were no terms considered essential in 1966 that were not also considered essential in 1967. There were actually 281 terms considered essential in 1967. Of these, 261 match with 1966 essential terms leaving 20 which were considered essential in 1967 but not in 1966. Therefore, Cell D must be 20. Accounting for the remaining terms, there were 598 total terms in 1966; 261 were essential leaving 337 nonessential. However, of these 337 nonessential terms, 20 were considered essential in 1967 and are located in Cell D leaving 317 terms. In 1967 there were 598 terms of which 281 were essential. The essential terms are accounted for in Cells B and D. There were no nonessential terms from 1967 considered essential in 1966 as noted earlier; this condition is indicated in Cell A as 0. The remaining terms are nonessential terms and are calculated as 598 - 281 = 317. This figure matches the figure of nonessential terms from 1966 and is the value for Cell C.

		1967 Non Ess.	1967 Ess.	
1966	Ess.	(A) 0	(B) 261	261
	Non Ess.	(C) 317	(D) 20	337
		317	281	598

$$\phi_{max} = \frac{BC - AD}{\sqrt{(A+B)\,(C+D)\,(A+C)\,(B+D)}} = \frac{96831 - 0}{\sqrt{(261)\,(337)\,(317)\,(281)}}$$

$$\phi_{max} = 0.934$$

Observed degree of association of essential terms - interindexer consistency

If the concept of essential terms is of value, there must be some degree of consistency between different individuals as well as with the same individual at different times. Applying the same assumptions as for determining intraindexer consistency, the interindexer ϕ coefficient was calculated for two experienced indexers.

Observed data

LCH (1967)

FLS (1967)		Non Ess.	Ess.	
	Ess.	(A) 78	(B) 203	281
	Non Ess.	(C) 249	(D) 160	409
		327	363	690

$$\phi = \frac{BC - AD}{\sqrt{(A+B)(C+D)(A+C)(B+D)}} = \frac{50547 - 12480}{\sqrt{(281)(409)(327)(363)}}$$

$$\phi = .327$$

$$X^2 = N\phi^2$$

$$= 690\,(.327)^2$$

$$= 73.8$$

This value of X^2 is significant at $p > .999$.

Maximum possible degree of association of essential terms

The maximum possible degree of association was determined for interindexer consistency in the same way as for intraindexer consistency which was described previously.

LCH 1967

FLS 1967		Non Ess.	Ess.	
	Ess.	(A) 0	(B) 281	281
	Non Ess.	(C) 327	(D) 82	409
		327	363	690

$$\phi = \frac{BD - AC}{\sqrt{(A+B)\,(C+D)\,(A+C)\,(B+D)}} \qquad \frac{94887 - 0}{\sqrt{(281)\,(409)\,(327)\,(363)}}$$

$$\phi_{max} = 0.814$$

Comparing ϕ coefficients:

ϕ intraindexer $= 0.461$ $\qquad \phi_{max} = 0.934$

ϕ interindexer $= 0.327$ $\qquad \phi_{max} = 0.814$

Therefore, the degree of association between essential terms is significant regarding both intraindexer and interindexer reliability, although intraindexer reliability is better. The values are shown graphically in Figure 1.

APPENDIX III

TYPICAL INDEX CARDS FOR ALL INDEXERS INDEXING THE SAME DOCUMENT (Ref. 5) (29,748)

TITLE: Uniform Vacuum Ultraviolet Reflecting Coatings on Large Surfaces

ABSTRACT: A method for producing uniform vacuum-deposited films on large substrates from a single, small area evaporation source is described. The procedure is being applied to the deposition of a uniform coating of magnesium fluoride on a 39-inch diameter vacuum-deposited aluminum mirror for use in the far ultraviolet. The magnesium fluoride restricts the oxidation of the aluminum and a uniformly high reflectance surface is obtained. Results of tests on a large number of two-inch square samples, arranged to duplicate the size and paraboloidal shape of the mirror, indicate that a reflectance of 82 $\pm$ 2 percent at Lyman alpha can be expected.

FLS (1966)	FLS (1967)	LCH (1967)
***Reflective coatings	*** Reflective coatings	***Vapor deposition
Vacuum	***Reflectance	***Reflective coatings
*** Magnesium fluoride	***Ultraviolet radiation	***Magnesium fluoride
***Aluminum	***Vapor deposition	***Aluminum
Coating	Films	Vapor phase
***Mirrors	***Magnesium fluoride	*Vacuum chambers
***Ultraviolet radiation	***Aluminum	*Films
***Reflectance	***Mirrors	Size effect
Measurement	Aluminum coatings	*Apparatus
***Vapor deposition	Wavelength	Evaporation
Surfaces	Protective coatings	*Aluminum coatings
Films	Coating	***Reflectance
	Manufacturing technology	Measurement
	Surfaces	***Mirrors
		Flow
		***Ultraviolet radiation
		Evaporators
		Design
		Theoretical analysis
		Coating

CODE:

* Term judged essential in one set of index terms
** Term judged essential in any two sets of index terms
*** Term judged essential in any three sets of index terms

PAT (1967)

Magnesium fluoride
Vapor deposition
Aluminum
Mirrors
Antioxidants
Reflective coatings
Reflectance
Optical properties
Ultraviolet radiation
Vacuum
Wavelength
Design
Coating
Measurement
Fabrication
Theoretical analysis
Central systems
Testing methods
Apparatus
Optical instruments
Surfaces

EJF (1967)

Reflective coatings
Reflectance
Vacuum
Ultraviolet radiation
Aluminum coatings
Vapor deposition
Films
Testing methods
Magnesium fluoride
Mirrors
Theoretical analysis
Fabrication
Protective coatings

LCS (1966)

Magnesium fluoride
Protective coatings
Reflectance
Ultraviolet radiation
Aluminum
Vapor deposition
Aluminum coatings
Space vehicles
Reflective coatings

TRD (1966)

Vapor deposition
Films
Magnesium fluoride
Mirrors
Aluminum coatings
Ultraviolet radiation
Space vehicles
Reflectance
Reflective coatings

TABLE 5. FREQUENCY DISTRIBUTION OF INDEX TERMS

TERM	No. of Indexers Using the Term
Magnesium fluoride	7
Reflectance	7
Reflective coatings	7
Ultraviolet radiation	7
Vapor deposition	7
Mirrors	6
Aluminum	5
Aluminum coatings	5
Coating	5
Films	5
Protective coatings	4
Measurement	3
Surfaces	3
Theoretical analysis	3
Vacuum	3
Apparatus	2
Design	2
Fabrication	2
Space vehicles	2

TABLE 5. (Continued)

TERM	No. of Indexers Using the Term
Testing methods	2
Wavelength	2
Antioxidants	1
Control systems	1
Evaporation	1
Evaporators	1
Flow	1
Manufacturing Technology	1
Optical instruments	1
Optical properties	1
Size effect	1
Vacuum chambers	1
Vapor phase	1

APPENDIX IV

SEARCH REQUESTS PROCESSED
December 1966 to December 1967

Search No.	Requestor	Organization	Search Title
887	M. D. Ward	Hercules Powder Co. Utah	Failure Criteria for Alloys
888	STINFO	WPAFB	Hydr. Lubr. System Contaminating
889	Mr. Burginas	Lawrence Radiation Lab.	Strength Degradation of Fil. wd. Comp. Long Time load
890	J. Sullivan	WPAFB	Hydrol. Stab. of Siloxane or Silicate Hydraul Fl.
891	F. Meyer	WPAFB	Corrosion of Ti in Organic Acids
892	STINFO	WPAFB	Chromium Alloy for Turbine Blades
893	Lt. Carey	WPAFB	Radiation from Projectile Impact
894	F. Meyer	WPAFB	Protective Coatings for Superalloy
895	F. Meyer	WPAFB	Corrosion of Metal Composites
896	E. J. Erwood	Xanthippe Chemicals	Electrically Conduct. Greases
897	D. Wetzler	AFC Industries	Titanium-Al-V Effect on Iron on Properties
898	D. Wetzler	AFC Industries	Springs of Titanium-Al-Cr-Fe-Mo
899	H. Smith	Boeing Airplane	Mechanical Properties of Germanium
900	J. Charlesworth	WPAFB	Measurement Testing of Thermal Properties
901	F. Scheffler	UD	Organic Semiconductors
902	K. S. Mazdigasni	WPAFB	Decomposition of Chelates
903	R. Hanschka	Hamilton Standard	Fatigue, Creep, Tensile Prop. on Aluminum 6061
904	G. A. Gegel	WPAFB	Atomization of Liquids, Metals
905	R. Wood	Aerojet General	Oxidation Protective Coatings
906	Dr. A. Prince	WPAFB	Meas. of Electrical Conductivity of Solutions
907	R. G. Prince	Straza Industries	Optical Properties of Invar.
908	M. Manoff	WPAFB	Speed of Sound in Plastics
909	L. Phillips	WPAFB	Fatigue of Glass, Boron, Reinforced Plastics
910	A. Yaross	WPAFB	Cost Compar for Fabric Ti, Ti Alloys
911	Mr. Stanley	Martin Orlando	Thermal Prop. of Phenoxy, Plastic Foams
912	W. Berner	UD	Rocket Nozzle Material Compatibility with Exhaust Gas
913	R. Racca	Nichols Engineer Inc.	Tuned Viscoelastic Dampers

Search No.	Requestor	Organization	Search Title
914	F. Scheffler	UD	System Updating-Cleaning Compounds
915	W. Scardino	WPAFB	Surface Preparation for Adhesive Bonding
916	F. Scheffler	UD	System Updating-Chemical Compounds
917	B. Graves	UD	Fluorine and Liquid Fluorine
918	R. Hanschka	Hamilton Standard	Size, Hardness Correlation with Mechanical Properties
919	H. Schwartz	WPAFB	Vacuum Effect on Polyimide Coating Film
920	A. F. Dahl	Brunswick Corporation	Fabrication of Polyimide Adhes. Bond Honeycomb
921	C. Lombard	WPAFB	Oxidation, Ductility of Fe, Co, Ni Alloys
922	F. Scheffler	UD	Update Elastomers
923	F. Scheffler	UD	Update Mylor
924	R. Greeson	Beech A/C Co.	Properties of Inconel 718 Sheets
925	R. Greeson	Beech A/C Co.	Properties of Cryoformed St. Steel 301
926	H. Roelefs	Interlake Steel Corp.	Friction Testing Package Bands
927	S. Mersal	WPAFB	High Temperature Insulation
928	Lt. H. Lachmann	WPAFB	R-541 Fatigue-Creep Interaction
929	L. H. Schwartz	ManLabs	Aircraft Armor
930	H. Doyle	Boeing Vertol	Strength Properties of Comp. Material
931	J. Crosby	WPAFB	Long Term Creep of Superalloys
932	H. Doyle	Boeing Vertol	Strength of Composite Components
933	M. Karnowsky	Sandia Corp.	Stress Strain of Metals-Curves
934	Dr. J. G. Faller	Boeing Vertol	Fracture Surfaces of Nonmetallic Elec. Microscope
935	R. Lawson	EG&G Inc.	Pool Boiling Heat Transfer for Clear Liquids
936	Dr. Jukkola	WPAFB	Compatibility of Structural Mat. with Liquid Metals
937	F. Meyer	WPAFB	Liquid Fluorine Effect on Dynamic Properties
938	L. C. Immel	Arthur D. Little	Asbestos Applications
939	Capt. D. R. James	WPAFB	Carbon Reinforced Metal Composite Interaction
940	Capt. D. R. James	WPAFB	Sonic Test of Metals
941	Capt. D. R. James	WPAFB	Fracture of Body Centered Cubic Metals Mech.

Search No.	Requestor	Organization	Search Title
942	Capt. D. R. James	WPAFB	Hot Ductility of Refractory Alloys
943	Capt. D. R. James	WPAFB	Superalloy Single Crystals
944	Mr. Booth	Turco Products	Increasing Titanium Alloys - Information Since 1962
945	G. Young	WPAFB	Properties of Potassium
946	F. Meyer	WPAFB	Diffusion Coatings
947	H. O. Rennat	WPAFB	Fabrication of Metal Composite
948	F. McClintock	MIT	Fracture Toughness of High Strength Steels
949	R. Hanschka	Hamilton Standard	Properties of Carborized Steel
950	R. Hanschka	Hamilton Standard	Properties of Cast Inconel 718
951	S. Zakangez	WPAFB	Heat of Fusion, Specific Heat of Lithium
952	D. Stevison	WPAFB	Thermal Diffusivity of Hafnium
953	W. Conrardy	WPAFB	Hydrogen Embrittlement of Titanium Alloys
954	B. Emrich	WPAFB	Ballistic Properties of Ceramics
955	R. Bowman	WPAFB	Joining Metal Composite Structure
956	W. Conrardy	WPAFB	Hydride Formation in Titanium, Ti Alloys
957	Dr. Strnat	WPAFB	Magnetic Prop, Crystal Structure of Rare Earth Cobalt System
958	J. Webb	WPAFB	Density of Solid Methane
959	F. Meyer	WPAFB	Corrosion of Structural Mat. by Liq. Propellants
960	Capt. D. R. James	WPAFB	MS Rel Mech Prop. Ni-Alloys
961	Capt. D. R. James	WPAFB	Physical Metallurgy of Nickel Alloys
962	Dr. H. Rosenburg	WPAFB	Photochromic Materials
963	C. Slenk	Carborundum Co.	Fibrous Thermal Insulation
964	Mrs. Chapman	Raytheon	Dispersion Hardened Platinum Rh
965	W. R. McIntyre	MTO Tinker	Cleaning Jet Turbine Blades
966	J. W. Pitts	WPAFB	Solvent Pollution of Atmosphere
967	R. Bidula	WPAFB	Environment Effects on Ti, Alloys
968	J. Payer	WPAFB	Properties of Pure Iron
969	J. Kasten	WPAFB	Metallurgy of Titanium, Ti Alloys
970	J. Kasten	WPAFB	Textured Titanium
971	H. S. Schwartz	WPAFB	Reinforced Plastics Used in Aircraft
972	T. Reinhart	WPAFB	Adhesive Bonding of Laminates, Etc.
973	T. Reinhart	WPAFB	Impact Resist. Adh. Bonded Structures
974	W. Reiman	WPAFB	Mech. Prop. of Ti-Al-V Forgings
975	W. Reiman	WPAFB	Cancelled on Request

Search No.	Requestor	Organization	Search Title
976	W. Reiman	WPAFB	Cancelled on Request
977	Mr. Canter	WPAFB	Pulse Power Amplifiers/Space Applications
978	F. Meyer	WPAFB	Protective Coatings for Be, Be Alloys
979	G. H. Hicheiman	WPAFB	Protective Coatings for Turbine Compressor Comp.
980	J. Wurst	UD	Rain Erosion Resistance
981	J. Wurst	UD	Analytical Design, Eval. of Rocket Noz. Inserts
982	E. W. McKelvey	WPAFB	Nondestructive Testing Techniques
983	C. Harmsworth	WPAFB	Mechanical Metallurgical Prop. of B-Ti Forgings
984	Capt. W. Schulz	WPAFB	Boron Reinf. Aluminum Matrix Composites
985	J. Ferguson	WPAFB	Numerically Controlled Machining
986	R. Hule	WPAFB	Standard Combustion Tests
987	Lt. C. W. Silha	WPAFB	Deformation Mechanism of Binary Alloys
988	S. Davis	WPAFB	Cathodic Corrosion Protection
989	J. Wurst	UD	High Press Plasma Arc Design
990	Mrs. Jajudi	Olin Research Center	Thermally Stable Polymers Research
991	T. Reinhart	WPAFB	Aircraft Tires Testing and Technology
992	H. Mittendorp	WPAFB	Inclusion Effect on Steel Properties
993	H. Mittendorp	WPAFB	Nondestructive Testing for Inclusion in Steel
994	H. Mittendorp	WPAFB	Melting Effect on Steels
995	M. Wexler	General Electric MSD	Rigid Organic Foam Insulation
996	M. Wexler	General Electric MSD	Carbon Foams-Properties
997	D. Stevison	WPAFB	Optical Properties of Beryllia
998	P. Smoot	Army Material Research	Toughness Titanium-Al-Sn-V
999	P. Smoot	Army Material Research	Fatigue Crack Titanium-Al-Sn-V
1000	P. Smoot	Army Material Research	Thermal Proc. Titan-Al-Sn-V
1001	R. A. C. Fearn	Babcock & Wilcox	Effect of Hot Water on Bearing
1002	T. Membry	Singer Co.	Coef. of Friction for Plastics
1003	L. Schwartz	Esso Research	Request Cancelled
1004	J. H. Ross	WPAFB	Fibrous Decelerator Materials
1005	J. H. Ross	WPAFB	Carbon Fibers from Extruded Polymers
1006	Lt. Chronister	WPAFB	Thermo Prop. of Noble Gases at Cryog. Temp.

Search No.	Requestor	Organization	Search Title
1007	Lt. Chronister	WPAFB	Cryogenic Properties of Metals
1008	Lt. Chronister	WPAFB	Properties of Intermetallics
1009	C. S. Cook	WPAFB	Forging Evaluation of Ti Alloys
1010	C. S. Cook	WPAFB	Spin Test Ti Alloys
1011	E. Miller	WPAFB	Man. of Elec. Circ. W. Si. Nitride
1012	H. Lisker	Avco Corp.	Corrosion of Anhydrous Ammonia on Al or Ti
1013	W. Johnson	WPAFB	High Temperature Polymer Insulation
1014	J. Sieron	WPAFB	High Temperature Polymer Fillers
1015	I. Holehouse	WPAFB	Flexural Fatigue Data
1016	J. H. Muntz	WPAFB	Sol Vap for Atomic Absorp Spectra
1017	P. M. Hemenger	WPAFB	II-IV Semiconductor Compounds
1018	R. W. Dimiduk	WPAFB	Dev. of High Pressure Mass Spectro. Probe
1019	W. Frederick	WPAFB	Research Hi Power Eff. Ferrites
1020	M. T. Ryan	WPAFB	Enhance Nuclear Magnetic Resonance Spec.
1021	V. L. Donlan	WPAFB	Grow & Eval. Trans. Met. Sen. Rel.
1022	C. M. Phillipi	WPAFB	Fast Scan Interferometer
1023	W. Frederick	WPAFB	Pyro Bn for Electromagnetic Windows
1024	W. Lawson	Contractor	Fracture Mech. & Crack Propaga. Since 1966
1025	Capt. Horsfield	WPAFB	Failure of MW Tubes
1026	G. L. Campbell	WPAFB	Rolling Taper Plates
1027	T. S. Felker	WPAFB	Fatigue Resistant Titanium Alloy Extrusions
1028	G. L. Campbell	WPAFB	Explosive Bonding Techniques
1029	G. L. Campbell	WPAFB	Prod. of Panels with Cooling Ducts
1030	G. L. Campbell	WPAFB	Prod. of Wheels, Discs of Hi Temp. Alloys
1031	Mr. Poliquin	WPAFB	Packaging S-S Circuits
1032	G. A. Gegel	WPAFB	Prod. of Ti or Ti Alloy Powders
1033	G. E. Eichelman	WPAFB	Suppression of Infrared
1034	Mr. Bratt	WPAFB	C owth of Si Single Crystals
1035	A. S. Dalton	WPAFB	Vapor Deposition of Metal Coatings
1036	E. Miller	WPAFB	Growth of Si Single Crystals on Sapphire
1037	Capt. Horsfield	WPAFB	Technical Joining Waveguides
1038	E. Miller	WPAFB	Mechanical Technology of Semiconductors
1039	L. C. Polley	WPAFB	Impregnated Tungsten Rocket Nozzles
1040	G. A. Gegel	WPAFB	Explosive Form PWD Metals
1041	L. C. Polley	WPAFB	High Energy Forging Apparatus

Search No.	Requestor	Organization	Search Title
1042	Lt. O'Hara	WPAFB	Textured Titanium Sheet
1043	Lt. O'Hara	WPAFB	Titanium Alloy Ingot Purification
1044	K. L. Love	WPAFB	Superalloy or Refractory Alloy Tubing
1045	Mr. Trinkle	WPAFB	Ferrite Delay Lines
1046	Mr. Bailer	WPAFB	Manufacturing Processes for Memory Sys.
1047	Mr. Bratt	WPAFB	Hydroth. Growth of BeO Sin Crystals
1048	K. L. Love	WPAFB	Ceramic/Metal Composite Armor
1049	Mr. Poliquin	WPAFB	Joining S-S Circuits
1050	C. T. Fouse	WPAFB	Emission Spectra at High Temperatures
1051	E. Morrissey	WPAFB	Flame Retardant Polymers
1052	W. Rooney	WPAFB	Flammability in Oxygen-Rich Atmosphere
1053	F. L. Whitney	WPAFB	Forming Metals Ultsound
1054	Capt. G. Purcell	WPAFB	Thermal Optical Properties of Niobium
1055	S. Carson	WPAFB	Triplet-Triplet AB Spectra
1056	Dr. H. L. Gegel	WPAFB	Stress Corrosion of Titanium, Al, Steels
1057	Mr. Campbell	WPAFB	Magnetic Prop. of Metal-Sulfur Complexes
1058	P. W. Dimiduk	WPAFB	Structural Degradation of Kryptonated Metals
1059	W. R. McIntyre	WPAFB	Electrostatic Filters
1060	E. J. Rolinski	WPAFB	Thermal Transport of Gases
1061	V. L. Donlan	WPAFB	Wave Functions of Rare Earth Ions
1062	A. K. Hopkins	WPAFB	Shock Wave Propagation
1063	V. L. Donlan	WPAFB	Crys. Growth of Sr-Ba Ti Solid Sol Crystals
1064	B. Emrich	WPAFB	Mat. Testing in Vacuum at Elevated Temps.
1065	E. Abshire	Albany Metallurgical Research Center	Tungsten-Dispersion Str. & Ele. Temp. Prop.
1066	J. B. Christian	WPAFB	Thixotropic Greases
1067	T. J. Aponyi	WPAFB	Development of Poly. Struc. Adhesives
1068	S. Litvak	WPAFB	High Temp. Polymer Adhesive Bonding
1069	L. Kopell	WPAFB	Pyrolytic Boron Nitride
1070	Mr. Trinkle	WPAFB	Manufacturing Tech TV Camera Tubes
1071	C. W. Caster	WPAFB	Manufacturing Technology Amplifiers
1072	E. S. Smith	WPAFB	Residual Stress Measurement
1073	G. W. Trickett	WPAFB	Titanium Wires and Cables
1074	W. T. O'Hara	WPAFB	Nondestructive Test Casting
1075	C. Tanis	WPAFB	High Modulus Filament Wound Plas. Compo.

Search No.	Requestor	Organization	Search Title
1076	C. Tanis	WPAFB	Glass Film Reinforced Plastics
1077	S. Litvak	WPAFB	Manu. Filament Wound Plastic Structures
1078	L. J. Conlon	WPAFB	Elastomer Monomer Production
1079	R. L. Kennard	WPAFB	Production of Intermetallic Compounds
1080	C. S. Anderson	WPAFB	Manufacturing Plasticizers
1081	D. L. Schmidt	WPAFB	Attaching Ablative Heat Shields
1082	~~D. L. Schmidt~~	~~WPAFB~~	New Ablative Composites for Rocket Nozzles
1083	D. L. Schmidt	WPAFB	3-D Reinforced Ablat Plastic Comp.
1084	D. L. Schmidt	WPAFB	Microwave Transparent Ablative Polymers
1085	D. L. Schmidt	WPAFB	Theoretical Concepts for Eros. Res. Abl. Comp.
1086	D. L. Schmidt	WPAFB	Composite Thermal Prot. Sys. for Nuclear Explosions
1087	T. J. Aponyi	WPAFB	Surf. Preparation, Primers for Met. Adhesives
1088	J. D. Ray	WPAFB	High Temperature Ceramic Matrix Composites
1089	J. D. Ray	WPAFB	Microdynamic Behavior of Comp.
1090	J. D. Ray	WPAFB	Graphite Fiber Surface Treatment of Composites
1091	Capt. Russo	WPAFB	Impact Strength of Large Steel Sections
1092	F. Scheffler	WPAFB	Update S-glass - Paula
1093	E. Morrissey	WPAFB	Heat Resistant Polymers
1094	W. L. Rooney	WPAFB	Combustion in an Oxygen Rich Atmosphere
1095	Packard-Hewlett	WPAFB	Diffusion Coefficient of Copper
1096	Lt. C. W. Canter	WPAFB	Rapid Heating Electron Tube Cathodes
1097	B. D. McConnell	WPAFB	Properties of Dry Film Lubricant
1098	F. H. Bair	WPAFB	Survey of High Temperature Resins
1099	E. Starr	Contractor	Metal Properties at Cryogenic Temps.
1100	J. M. Pasick	WPAFB	Electrochemical Machining
1101	R. M. Van Vliet	WPAFB	Fiber Optics
1102	B. Ruben	WPAFB	Fatigue Testing of Metal Honeycomb
1103	W. H. Wheeler	WPAFB	Design of Brittle Aerospace Structures
1104	R. H. Udlock	General Dynamics	Taber Abrasion Tests - 140 Abrasor
1105	R. Watson	Iowa State University	Thermal Conductivity of Liquid
1106	C. Harmsworth	WPAFB	Properties of Aluminum 7075
1107	J. M. Pasick	WPAFB	Integrity of Metal Surfaces
1108	Mr. Shoemaker	Coats & Clark	Ultrasonic Sintering of Powdered Metals

Search No.	Requestor	Organization	Search Title
1109	A. C. Meeks	WPAFB	Aromatic Polyamides
1110	K. Herman	Airesearch Mfg. Co. Arizona	Elect. Resistivity & Magnet. Perme. of WC & TiC
1111	R. Watson	Iowa State University	Thermal Conduct. of Non Newtonian Liquids, Suspens.
1112	R. Askins	UD	Temperature Effect on Stress-Strain Work
1113	Capt. D. R. James	WPAFB	Metallography of TD-Ni
1114	Capt. D. R. James	WPAFB	Crystal Growth of TD-Nickel
1115	Capt. D. R. James	WPAFB	TD Nickel Hard Mechanics
1116	R. Askins	UD	Load Deflection of Composite Material
1117	R. A. Heaver	Racine & Vickers Armstrong	Mech. Prop. of Composites at Low Temperature
1118	R. F. Wilkinson	Clary Corp.	Epoxy Coating on Miniature Bearings
1119	J. Pasick	WPAFB	Chemical Milling Technology
1120	J. Pasick	WPAFB	Holography
1121	M. I. Greene	American Cyanamid Company	Polymers for Desalination Application
1122	E. Cutrell	Martin Denver	Bending Fatigue Stainless 301
1123	V. DePierre	WPAFB	Specific Heat of U-700
1124	A. Heideman	Lockheed Cal.	Mechanical Prop. of D6AC
1125	A. Heideman	Lockheed Cal.	Comp. of Ti, Ti Alloys with Fuel, etc.
1126	E. Abshire	Albany Metallurgy Research Center	Dispersion Hard Tungsten
1127	J. Weidner	UD	Curing Temperature Effect on Compos.
1128	S. Bradstreet	Dayton, Ohio	Composition Effect on Crys. Struc. in Alk. Hal.
1129	R. Henderson	WPAFB	Electron Diffraction & Microscopy
1130	E. Horne	WPAFB	Fibers for Expandable Structures
1131	G. Williams	Bridgeport Brass Co.	Ceramic Mat. for Extrusion Drawing of Metals
1132	S. Zakanycz	WPAFB	Liquid Metal Heat Exchangers
1133	S. Zakanycz	WPAFB	Viscosimetry-Apparatus & Technology
1134	G. E. Metzger	WPAFB	Joining Technology at WPAFB
1135	W. Mahieu	UD	Tensile Creep of Composites
1136	J. H. Schwartz	Olin Research Center	Polymers Since 1967
1137	F. Ostermann	WPAFB	Fatigue & Heat Treatment of Al-7075-T73
1138	J. Stock	Fairchild Hiller	Mech. Prop., Toughness of Ti-Al-Sn-V
1139	C. A. Lombard	WPAFB	Oxidation, Ductility, Str of Ti Alloys at Hi Temps.

Search No.	Requestor	Organization	Search Title
1140	P. Thomas	UD	Infrared Spectroscopy of Carboxylic Acid Salts
1141	F. Fruin	UD	Boron Hydrides Anion Molecular Structure
1142	Lt. Bass	WPAFB	Fluidic Controls for Gas Turbines
1143	A. Heideman	Lockheed Calif. Co.	Fatigue of Steels and Ti Alloys
1144	A. Heideman	Lockheed Calif. Co.	Fatigue Data for Steels, Ti Alloys
1145	C. Bolenca	UD	Graphite Fiber Surface Treatment
1146	C. Bolenca	UD	Surface Treatment of Fibers for Composites
1147	F. Reinhart	WPAFB	Coatings or Bands for Projectiles
1148	S. Zakanycz	WPAFB	Mass Diffusivity of Helium and Air
1149	J. Charlesworth	WPAFB	Melting Points of Non Metals
1150	J. Charlesworth	WPAFB	Melting Point of Metals & Alloys
1151	J.T. O'Brien	WPAFB	Flammability of Nonmetallic Fibers
1152	R. Dominic	UD	Fatigue of Al-7075-T6
1153	S. L. Hanlein	US Naval Ord. Lab.	Fracture and Stress of Glass
1154	S. L. Hanlein	US Naval Ord. Lab.	Handbooks & Symposia for Acoustical Prop.
1155	USAF HQs.	USAF HQ	Sierracoat RS Antireflective Coatings
1156	M. Guenther	WPAFB	Hydrostatic Forming
1157	W. Jater	WPAFB	Sublimation of Brasses
1158	T. Dolce	WPAFB	Phenolic-Refrasil Decomposition
1159	F. E. Frost	Lawrence Radiation Laboratory	Mechanical Prop. and Behavior of Foams
1160	P. Monaghan	WPAFB	Catalysts for Synthesis of CO_2
1161	P. Monaghan	WPAFB	Cathode Sputtering
1162	L. Parsons	WPAFB	Fatigue of Inconel 718
1163	W. Scardino	WPAFB	Bonding or Repair of Metal Plastic Struc.
1164	F. Scheffler	WPAFB	Composites and Laminates
1165	E. Pieracy	Howmet Corp.	Diffusion Coatings for Superalloys
1166	G. E. Metzger	WPAFB	Joining of Titanium
1167	A. Olevitch	WPAFB	Outgassing in Vacuum or Oxygen
1168	A. Olevitch	WPAFB	Fire Retarding Coating or Plastics
1169	J. Rhodehamel	WPAFB	Low Temp. Curing Polymer Matrix
1170	E. Matson	Goerz Optical Co.	Splicing of Mylar and Estar
1172	R. Spain	WPAFB	Mechanical Properties of Porous Polymeric Fibers
M159	Lt. Peyton	WPAFB	Compatibility of H_2O_2 with Zinc Pow. Paints

REQUESTOR INDEX

Ferguson, J. 985	APFG-36	AFAPL
Fouse, G. T. 1050	AVP	AFAL
Frederick, W. G. D. 1019, 1023	MAYT	AFML
Frost, F. E. 1159	Lawrence Radiation Lab.	
Fruin, E. 1141	UDRI	
Gegel, G. A. 904, 1032, 1040	MATB	AFML
Gegel, Dr. H. L. 1056	MAMS	AFML
Graves, B. 917	UDRI	
Greene, M. I. 1121	American Cyanamid Co.	
Greeson, R. 924, 925	Beech a/c Co.	
Guenther, M. A. 1156	MATF	AFML
Hanlein, S. L. 1153, 1154	US NOL	
Hanschka, R. 903, 918, 949, 950	Hamilton Standard	
Harmsworth, C. L. 983, 1106	MAAM	AFML
Heaver, R. A. 1117	Racine and Vickers Armstrong, Inc.	
Heideman, A. 1124, 1125, 1143, 1144	Lockheed California Co.	

Hemenger, P. M. 1017	MAYE	AFML
Henderson, R. 1129	MAAS	AFML
Herman, R. 1110	Airesearch Mfg. Co.	
Hichelman, G. H. 979	MATE	AFML
Holehouse, I. 1015	MAAM	AFML
Hopkins, A. K. 1062	MAYH	AFML
Horne, E. 1130	MAAM	AFML
Horsfield, Capt. 1025, 1037	MATE	AFML
Huie, R. 986	APFT-30	AFAPL
Immel, L. C. 938	Arthur D. Little, Inc.	
Jajudi, Mrs. 990	Olin Research Center	
James, Capt. D. R. 939, 940, 941, 942, 943, 960, 961, 1113, 1114, 1115	MAAE	AFML
Jater, W. 1157		
Johnson, W. P. 1013	MATC	AFML
Jukkola, Dr. 936	MAAA	AFML
Karnowsky, M. M. 933	Sandia Corp.	

Kasten, J. 969, 970	MAMS	AFML
Kennard, R. L. 1079	MATB	AFML
Kopell, L. 1069	MATC	AFML
Lachmann, Lt. H. 928	MAAM	AFML
Larson, R. 935	EG&G Inc.	
Lawson, W. 1024	Pompano Beach	
Lisker, H. 1012	Avco Corp.	
Litvak, S. 1068, 1077	MATC	AFML
Lombard, C. A. 921, 1139	MAMP	AFML
Love, K. L. 1044, 1048	MATB	AFML
Mahieu, W. 1135	UDRI	
Manoff, M. 908	MANC	AFML
Matson, E. 1170	Goerz Optical Co.	
Mazdigasni, K. S. 902	MAMC	AFML
McClintock, F. 948	MIT	
McConnell, B. D. 1097	MANL	AFML

[illegible]	[illegible]	AFML

[illegible] | [illegible] | AFML
1162 | |

[illegible], J. M. | MATF | AFML
1100, 1107, 1119, 1120 | |

Payer, J. | MAMS | AFML
966 | |

Peyton, Lt. K. S. | MAAS | AFML
M159 | |

Phillipi, C. M. | MAYA | AFML
1022 | |

Phillips, L. | AFFDL |
909 | |

Pieracy, E. | Howmet Corp. |
1165 | |

Pitts, Lt. Col. J. W. | MAAS | AFML
966 | |

Poliquin, P. | MATE | AFML
1031, 1049 | |

Polley, L. C. | MATB | AFML
1039, 1041 | |

Prince, Dr. A. | |
906 | |

Purcell, Capt. G. | MAAE | AFML
1054 | |

Racca, R. | Nichols Engineering Inc. |
913 | |

Ray, J. D. | MANC | AFML
1088, 1089, 1090 | |

Reiman, W. | MAMD | AFML
974, 975, 976 | |

Reinhart, T. | MAAE | AFML
972, 973, 991, 1147 | |

Schwartz, I. 1003	Esso Research
Schwartz, L. H. 929	ManLabs.
Shoemaker, Mr. 1108	Coats and Clark Inc.
Sieron, J. K. 1014	MANE AFML
Silha, Lt. C. W. 987	APFT-29 AFAPL
Slenk, C. 963	Carborundum Co.
Smith, E. S. 1072	MATF AFML
Smith, H. 899	Boeing Airplane Co.
Smith, R. G. 907	Straza Industries
Smoot, P. 998, 999, 1000	Army Material Research
Spain, R. 1172	MANC AFML
Stanley, Mr. 911	Martin/Orlando
Starr, E. 1099	Warren-Lawson
Steirson, D. 952, 997	MAAE AFML
Stock, J. 1138	Fairchild Hiller
Strnat, Dr. 957	MATE AFML

INDEX OF REQUESTING ORGANIZATIONS

Report for Period 1 December 1966 to 1 December 1967

Organization	No. of Searches
Aberdeen Proving Grounds	
Aerojet General	1
AFAL	1
AFAPL	3
AFC Industries	9
AFFDL	2
AFML	1
AFSC	175
Airesearch Mfg. Co.	5
Albany Metallurgy Research Center	1
American Cyanamid Co.	2
Army Material Research	1
Arthur D. Little, Inc.	3
Avco Corp.	1
Babcock and Wilcox, R&D Division	1
Beech a/c Co.	1
Boeing Airplane Co.	2
Bridgeport Brass Co.	4
Brunswick Corp	1
Carborundum Co.	1
Clary Corp.	1
Coats and Clark, Inc.	1
Dayton, Ohio	1
EG and G, Inc.	1
Esso Research	
Fairchild Hiller	1
General Dynamics	1
General Electric MSD	1
Goerz Optical Co.	2
Hamilton Standard	1
Headquarters USAF	4
Hercules Powder Co.	1
Howmet Corp.	1
Interstate Steel Corp.	1
Iowa State University	1
Lawrence Radiation Lab.	2
Lockheed California Co.	2
ManLabs.	4
Martin	1
MIT	2
	1

Organization	No. of Searches
Nichols Engineering, Inc.	1
Olin Research Center	2
Packard Hewlett	1
Pompano Beach	1
Racine and Vickers, Armstrong, Inc.	1
Raytheon	1
Sandia Corp.	1
Singer Co.	1
Strana Industries	1
Tinker AFB	2
Turco Products	1
University of Dayton	21
U. S. Naval Ordinance	2
Warren-Lawson	1
Xanthippe Chemicals	1
Unclassified	2

Total number of searches for this period: 286

APPENDIX V

SUBJECT CATEGORIES: DOCUMENT INPUT AND SEARCHES PROCESSED BY SUBJECT CATEGORY

AMIC	COSATI	CATEGORY
01	01	Aeronautics
		Aerodynamics
		Aeronautics
		Aircraft
		Aircraft flight control and instrumentation
		Jet engines
02	03+04	Astronomy, Astrophysics, Atmospheric Sciences
		Astronomy
		Astrophysics
		Atmospheric physics
		Meteorology
03	06+07	Chemistry, Biology, Medical Sciences
		Biochemistry
		Bioengineering
		Biology
		Chemical analysis
		Chemical engineering
		Inorganic chemistry
		Life support systems
		Organic chemistry
		Physical chemistry
		Radiochemistry
		Toxicology
04	09	Electronics and Electrical Engineering
		Components
		Electronic & electrical engineering
		Telemetry

AMIC	COSATI	CATEGORY
05	11A	Adhesives
		Ceramic cements
		Organic resin adhesives
		Potting compounds
06	11A	Seals, Sealants
		Ceramic-metal bonds
		Mechanical seals
		O-rings
07	11B	Ceramics, Refractories, Glasses, Minerals
		Borides
		Carbides
		Carbon, graphites
		Mixed oxides
		Nitrides
		Single oxides
08	11C	Coatings, Paints, Oxide Films
09	11D	Composites Materials, Laminates, Sandwich Structures, Honecomb
10	11E	Fibers, Textiles, Cloth
11	11F	Metallurgy, Metallography
		Alloys
		Metals
12	11H	Oils, Lubricants, Heat Transfer Fluids, Greases, Hydraulic Fluids
13	11I	Polymers, Plastics
14	11J	Elastomers
15	11K	Cleaning Compounds, Surface Active Agents

AMIC	COSATI	CATEGORY
16	11L	Wood and Paper Products
17	21	Fuels, Propellants, Propulsion Systems, Explosives
18	13	Mechanical, Industrial, Civil and Marine Engineering
		Civil engineering
		Contruction equipment, materials, supplies
		Containers and packaging
		Coupling, fittings, fasteners, joints
		Industrial processes
		Machining, tools, machine elements such as bearings, gas lubrication systems
		Marine engineering
		Pumps, filters, pipes, fittings, tubing, and valves
		Safety engineering
		Structural engineering
19	14	Methods and Equipment
		Apparatus
		Detectors
		Laboratories, test facilities, and test equipment
		Recording devices
20	18	Nuclear Science and Technology
		Fuel elements; fuel, nuclear
		Nuclear explosions
		Nuclear power plants
		Nuclear reactors
		Radiation shielding
		Radioactive wastes
21	20	Physics
		Acoustics

AMIC	COSATI	CATEGORY
21 (Continued)	20	Physics
		Crystallography
		Electricity and magnetism
		Fluid mechanics
		Masers and lasers
		Optics
		Particle accelerators
		Particle physics
		Plasma physics
		Quantum theory
		Solid mechanics
		Solid-state physics
		Spectrometry, spetroscopy
		Thermodynamics
		Wave propagation
22	10, 16, 22	Space Technology and Missiles
		Astronautics
		Energy conversion, solar cells
		Launch vehicles
		Missile technology
		Re-entry vehicles
		Rockets
		Satellites, artificial
		Spacecraft
		Trajectories and re-entry

TABLE 6. DOCUMENT INPUT AND SEARCHES PROCESSED BY SUBJECT CATEGORY

AMIC CATEGORY	DOCUMENT CATEGORIES		SEARCH CATEGORIES	
	No.	%	No.	%
01	139	2.9	6	1.8
02	38	0.8	1	0.3
03	361	7.5	17	5.0
04	137	2.9	18	5.3
05	20	0.4	8	2.4
06	35	0.7	0	0.0
07	170	3.5	20	5.9
08	116	2.4	12	3.6
09	134	2.8	29	8.6
10	30	0.6	4	1.2
11	973	20.3	97	28.7
12	352	7.3	6	1.8
13	200	4.2	22	6.5
14	43	0.9	4	1.2
15	2	0.0	2	0.6
16	2	0.0	0	0.0
17	54	1.1	1	0.3
18	363	7.6	31	9.2
19	404	8.4	14	4.1
20	173	3.6	1	0.3
21	902	18.8	32	9.5
22	144	3.0	13	3.8

APPENDIX VI

DESCRIPTION OF CAS SDI PROGRAM AND TYPICAL USER PROFILES

The SDI Experiment covers 32 journals of general chemical interest[2]; POST-J covers articles from approximately 425 journals and Government reports specifically dealing with polymers, and POST-P covers polymer related patents from 20 countries. POST-P and POST-J are subscribed to for members of the Polymer Branch of the Nonmetallic Materials Division (MAN) of the AFML.

Briefly, the operation of the SDI program involves text searching of the titles and abstracts in conjunction with user profiles by an IBM 360 Model 40 computer. The "hits" or retrievals are determined by those title-abstract units which qualify according to the user profile search logic. The hits are recorded by abstract number and bibliographical data for each user, and a printing of all abstracts processed in the particular period is provided. The user can then pick out the abstracts from the printing which correspond to the hits found for his particular profile.

The search program has a truncating feature by which portions of technical words or phrases can be employed in the user profile. The advantage of this truncating feature is that the root of a technical word can be listed and the text searching will then retrieve all variations of the root word as long as the root itself appears in the title or abstract. To illustrate, a scientist may be interested in all silane compounds. By listing *SILAN* as a search unit, document abstracts on silane, dimethylsilane, silanol, polysilanes, etc., would be retrieved. If he were interested in all silicon compounds, both organic and inorganic, the root *SIL* might be chosen. There is a possibility that noise or false retrievals may occur, e.g., the words silent, silk, tensile, and silver, could cause false retrieval in the preceding example, but the probability of false retrieval is very low with judicious selection of the truncated units. One can specify that the truncation must be the start of a word, i.e., preceded by one or more blanks by omitting the leading asterisk; likewise the truncation can be designated as the end of a word by leaving out the terminal asterisk, i.e., the truncation must be followed by one or more blanks.

The user search profiles are formulated by combining terms (including truncated terms) with AND, OR, NOT logic. In addition to the AND, OR, NOT logic, optional weighting factors can be applied either in the positive or negative direction. Thus undesirable terms can be given a negative bias without eliminating them altogether as would be the case with a strict NOT logic operator. With weighted terms one specifies a minimum acceptable weight which must be satisfied before retrieval occurs (after all other parameters are satisfied).

Thus in the case cited above of a man desiring information on silicon compounds and using *SIL* as a profile term, the terms TENSILE and SILVER could be given a negative weight. An abstract on tensile strength of stainless steels would thus not be retrieved, but an abstract on tensile elastic constants of single crystal silicon dioxide would satisfy the minimum weight requirements and thus be retrieved if the weights are carefully selected.[1]

The preparation of good user profiles is the key to the success of the SDI program. User profiles were established by a polymer chemist at AFML in conjunction with the users from the Polymer Branch. These profiles were reviewed by a consultant from CAS. Presently 12 profiles are being searched on a regular basis. Two of the profiles were modified considerably when it was indicated by user feedback that too much noise was occurring. The modifications of the profiles resulted in a considerable reduction in noise but very little, if any, loss of pertinent abstracts occurred. Modification in other profiles will undoubtedly be made as more experience is attained. Careful analysis of why the profile retrieved nonpertinent abstracts provides indications as to what modifications should be made to improve performance.

Typical examples of AFML search profiles currently in use are shown on the following pages.

SEARCH PROFILE FORM

REQUESTOR 075-08-01

AFFILIATION

ADDRESS

HEADER CARD [redacted]

OUTPUT INDICATORS	QUESTION WEIGHT

QUESTION: INFORMATION ON POLYMER DEGRADATION, THERMAL STABILITY AND THERMOGRAVIMETRIC ANALYSIS

DETAIL CARD

					LOGIC	TERM (DO NOT KEYPUNCH ASTERISK)	TERM WEIGHT
01	02	2	T	0	OR	POLY*	
01		3	T	0	OR	PLASTICS	
01		2	T	0	OR	ELASTOMER*	
01		2	T	0	OR	RESIN*	
01		2	T	0	OR	FIBER*	
01		2	T	0	OR	NYLON*	
01		2	T	0	OR	* AMIDE *	
02		2	T	0	OR	DEGRAD*	
02		2	T	0	OR	DECOMPOS*	
02		2	T	0	OR	PYROLYS*	
02		2	T	0	OR	DEPOLYMERIZ*	
02		2	T	0	OR	SCISSION*	
02		4	T	0	OR	* ZIP *	
02		3	T	0	OR	WEIGHT LOSS	
02		3	T	0	OR	WEIGHT CHANGE	
02		2	T	0	OR	THERMOGRAVIMETR*	
02		3	T	0	OR	TGA	
02		3	T	0	OR	DTA	
02		3	T	0	OR	THERMAL ANALYSIS	

SEARCH PROFILE FORM

REQUESTOR 075-10-01

AFFILIATION

ADDRESS

HEADER CARD

QUESTION NUMBER		OUTPUT INDICATORS	QUESTION WEIGHT
1		6 11	12 17
		1 ∅ ∅ ∅	

QUESTION INFORMATION ON THE PROPERTIES OF STIFF OR RIGID OR SEMI-FLEXIBLE POLYMERS

DETAIL CARD

DUPLICATE	PARA NO.	NO. PARA	WORD	TYPE	ROOT	LOGIC (DO NOT KP)	TERMS (DO NOT KEYPUNCH ASTERISK)	TERM WEIGHT
	∅ 1	∅ 3	4	T	0	OR	* STIFF *	
	∅ 1	DUPLICATE	4	T	0	OR	* RIGID *	
	∅ 1		4	T	0	OR	* FLEXIB *	
	∅ 2		2	T	0	OR	POLYMER*	
	∅ 2		2	T	0	OR	MACROMOLECULE*	
	∅ 2		2	T	0	OR	CHAIN*	
	∅ 3		3	T	0	OR	VISCOSITY	
	∅ 3		2	T	0	OR	SOLUB*	
	∅ 3		3	T	0	OR	SECOND VIRIAL COEFFICIENT	
	∅ 3		3	T	0	OR	LIGHT SCATTERING	
	∅ 3		2	T	0	OR	HYDRODYNAMIC*	
	∅ 3		2	T	0	OR	THERMODYNAMIC*	
	∅ 3		2	T	0	OR	MOLECULAR WEIGHT*	
	∅ 3		2	T	0	OR	VOLUME*	
	∅ 3		2	T	0	OR	PRECIPITATION*	
	∅ 3		2	T	0	OR	SEGMENT*	
	∅ 3		2	T	0	OR	CONFORMATION*	
	∅ 3		2	T	0	OR	CONFIGURATION*	
	∅ 3		2	T	0	OR	STATISTIC*	

SEARCH PROFILE FORM

REQUESTOR 075-11-01

AFFILIATION

ADDRESS

HEADER CARD				QUESTION: CARBORANE OR DECABORANE OR CARBOLLIDE
QUESTION NUMBER		OUTPUT INDICATORS	QUESTION WEIGHT	
1		1,∅,∅,∅		

DETAIL CARD

DUPLICATE	PARA. NO.	NO. OF PARA.	MODE	TYPE	LOGIC	LOGIC	TERMS (DO NOT KEYPUNCH ASTERISK)	TERM WEIGHT
	∅,1	∅,1	4	T	0	OR	* CARBORAN *	
	∅,1		4	T	0	OR	* CARBOLL *	
	∅,1		4	T	0	OR	* BARAN *	
	∅,1		2	T	0	OR	DECABORAN*	
	∅,1		4	T	0	OR	* CARBAMETALL *	

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) University of Dayton Research Institute Dayton, Ohio	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP

3. REPORT TITLE

INDEXER PERFORMANCE ANALYSIS AND OPERATIONS OF A DOCUMENT RETRIEVAL SYSTEM

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Final Summary Report, 1 December 1966 - 30 November 1967

5. AUTHOR(S) (Last name, first name, initial)

Scheffler, Frederic L.

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
February 1968	76	11

8a. CONTRACT OR GRANT NO. AF 33(615)-3389 b. PROJECT NO. 7381 c. Task No. 738103 d.	9a. ORIGINATOR'S REPORT NUMBER(S) 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) AFML-TR-67-379

10. AVAILABILITY/LIMITATION NOTICES

This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY AFML (MAAM) Wright-Patterson AFB, Ohio 45433

13. ABSTRACT

Two graduate students [illegible] chemistry were trained in indexing of technical documents through an e[illegible]ished training program. Certain modifications were made which tended to improve the training as evidenced by improved performance. A study of intra- and interindexer consistency was made by subjecting a group of technical documents to independent indexing by various individuals and by two of the same individuals one year later. The concept of essential index terms was introduced. Statistical analysis showed that intra- and interindexer correlation for experienced indexers of essential terms is significant with a probability $(p) > 0.999$. An evaluation of the indexer trainees showed that essential terms was a better performance criterion than total terms and that high interindexer consistency was achieved. Search requests declined from the preceding year, especially from the AFML. Possible reasons for a decline are presented. A Selective Dissemination of Information (SDI) Program with Chemical Abstracts was initiated. The development of a vocabulary and thesaurus for information science related documents has started. A free indexing technique is being used.

DD FORM 1 JAN 64 1473

UNCLASSIFIED
Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Air Force Materials Laboratory						
Indexing						
Statistical Analysis						
Personnel						
Evaluation						
Indexers						
Indexer Consistency						
Index Terms						
Thesaurus						
Document Retrieval Systems						
Information Retrieval Systems						
Searching						
SDI Program						
Information Science Documents						

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization *(corporate author)* issuing the report.

2a. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2b. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year, or month, year. If more than one date appears on the report, use date of publication.

7a. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7b. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8a. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8b, 8c, & 8d. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9a. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9b. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers *(either by the originator or by the sponsor)*, also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through
________________________________."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through
________________________________."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through
________________________________."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring *(paying for)* the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as *(TS)*, *(S)*, *(C)*, or *(U)*.

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, rules, and weights is optional.